Accessible Home Automation System for Elderly Individuals and People with Disabilities: Combining Eye Tracking, Hand Gestures, and Voice Recognition

TAKLA SASSINE, HASSAN MROUEH, ALEXANDRE YAACOUB, THALIA YAMMINE, ELIAS SAADEH, MOUSTAPHA EL HASSAN
Electrical Engineering Department, Faculty of Engineering
University of Balamand,
Kelhat,
LEBANON

Abstract: With increasing elderly populations and people with disabilities all over the world, there is an emerging need for accessible, autonomous home environments. This paper proposes a fully embedded, low-cost, and multimodal home automation system focused on solving interaction challenges for users with physical or sensory impairments. The proposed solution integrates eye-tracking, hand gesture recognition, and voice-based command inputs for a sustainably adaptable control towards diverse users. This system not only answers the needs of elderly and disabled people but also contributes to a global vision per the United Nations' Sustainable Development Goals. It was built using the Raspberry Pi 4 platform with open-source Python libraries, allowing the implementation of custom offline solutions with lower latency. Validation of the system was carried out by functionally deploying it at a caregiving facility, where it showed very high accuracy rates, from 90 to 96%, for the tested input modalities and robust performance under real-world conditions.

Key-Words: Home automation, hand gestures, eye tracking, speech interface, sustainability, assistive technology.

Received: May 19, 2025. Revised: July 13, 2025. Accepted: September 18, 2025. Published: May 22, 2026.

1 Introduction

With the global demographic shift toward an aging population coupled with increased disabilities, there is an intensified need for practical solutions to support independent living and enhance the quality of life among the affected people. More than 1 billion people, or roughly 15% of the world's population, have some kind of disability, which is set to grow even further due to aging populations and increased rates of chronic disease [1].

The integration of ICT into homes has developed significantly over the past twenty years [2], [3], [4]. The development of the X-10 home automation platform in 1975 allowed for the control of equipment through radio frequencies over distances, laying the base for more recent developments of smart home technology [5]. In the 1990s, more sophisticated systems came into being, such as IBM's Home Director software, which allowed users to take advantage of their personal computers to orchestrate a variety of subsystems within the home [6].

The concept of ambient intelligence has been evolving over the last few years; it describes those spaces that contain electronic devices with the ability to perceive and respond to the presence of a human, changing accordingly to allow better user interaction.

Despite these technological advances, the actual dissemination of home automation systems among elderly and disabled people has been slow. In a recent study investigating the determinants of adoption of smart home technology for aging-in-place, it emerged that while a high proportion of older people owned devices such as smartphones (80.8%) and tablets (74.3%), only a subgroup of them felt they were very competent in their use (46.3% and 42.7%, respectively), hence there is a need for even more intuitive and accessible solutions [7].

The advantages of home automation for people with disabilities are well-documented. A scoping review identified positive outcomes such as increased independence, improved mental health, improved social and community connectedness, improved safety, and support with activities of daily living [2], [8].

By 2050, people aged 65 and above will account for almost 23% of the Lebanese population, marking a clear demographic transition [9]. This aging is not unique to Lebanon, as the same trends are happening now worldwide, with the elderly population expected to be over 1.5 billion people by the middle of this century [10]. Simultaneously, an estimate of 10% to 15% of the Lebanese population in 2018 lives with some form of physical, sensory, intellectual, or mental disability [11]. For the majority of the population in these overlapping categories, such daily activities as the operation of domestic appliances are commonly reliant on external support.

Besides, the installed context-aware systems with the ability to understand natural user input like voice, hand, and eye gestures provide a new promise for independent living. As an example, eye-tracking can be used by a quadriplegic patient to open the blinds or turn the lights on/off [12], while a visually impaired user will be able to maintain environmental settings using speech-based interaction [13], [14]. Similarly, deaf or speech-disadvantaged users have benefited from gesture interfaces [15]. Therefore, the work proposes a complete home automation system based on eye-tracking, hand gesture recognition, and voice command interfaces, using the Raspberry Pi 4 platform and implemented with Python-based libraries. The system allows real-time low-latency control of home appliances, based on a user-centered design. Furthermore, the system is interfaced with face recognition for safe authentication so that access to the system is limited to certified users only [16].

This system was pilot-tested through field testing at the Wahat El Farah Association (El Koura, Lebanon), an organization that delivers services to older persons and persons with disabilities. Initial results indicated that the system demonstrated good usability, responsiveness, and flexibility across a range of impairments [17]. The aim of this work is two-fold: (1) to bridge existing accessibility gaps in home automation with the use of multiple input modalities and (2) to offer a sustainable, low-cost, scalable, and privacy-preserving assistive technology for independent living.

2 Problem Formulation

The field of smart home assistive technologies has evolved extremely rapidly over the last two decades, with the merging of Internet of Things (IoT) platforms, embedded systems, and multimodal human-machine interfaces. The main function of these systems is to enhance the independence and quality of life for individuals with physical, sensory, or cognitive impairments. In addition, the benefit of smart home technologies in alleviating caregiver burden and enabling "aging in place" is now established [18], [19].

Despite this momentum, commercially and academically advanced systems remain limited in scope, typically relying on a single modality for input (e.g., speech, gesture, or gaze). This severely constrains accessibility and accommodation, especially among users with compound impairments. Furthermore, penetration is low: only 19.3% of older adults report they have a voice assistant, and fewer than 6% use smart home hubs [20]. The primary reasons provided are usability, which is bad, high cost, and privacy [14], [21].

2.1 Eye-Tracking-Based Systems

Eye-tracking interfaces have been especially promising for those with impaired limb mobility. Mathew, Sreeshma, Jaison, Pradeep, and Jabarani demonstrated a home control system with Haar cascade classifiers and dlib-based landmark detection for blink and gaze control [12]. The system utilizes the Eye Aspect Ratio (EAR) and limbus tracking for detecting eye closures and gaze direction. Babu and Sivasubramaniam extended this line of inquiry with a low-cost webcam-based solution utilizing pupil position to drive the cursor through OpenCV [22]. Low-cost and relatively lightweight, the solution is marred by high latency and decreased accuracy under non-uniform lighting or background conditions. Importantly, these systems do not incorporate fallback modalities, rendering them less inclusive for individuals who encounter ocular instability or irregular blink patterns.

More recent platforms, such as the MAISON platform [3], demonstrate the growing trend for AI-driven, context-sensitive solutions, but most still focus on passive monitoring rather than interactive control. This gap demonstrates the need for eye-tracking systems that are components of active, multimodal control systems, which can adapt based on environmental and user limitations.

2.2 Multimodal Systems with Limited Scope

In an attempt to overcome the limitations of single-modality systems, researchers have explored dual-modality systems. Klaib, Alsrehin, Melhem, and Bashtawi proposed an IoT-based prototype combining eye tracking via Tobii and voice interaction via Amazon Alexa. The system accommodated 50 gaze events per second and facilitated easy home appliance control by gaze or voice [23]. While this two-modality approach made it more accessible, it still did not support gesture control and hence excluded users with speech impairments or users operating in noisy environments. Cloud-based voice recognition services and proprietary hardware (Tobii) also raise issues regarding data privacy, cost, and scalability in the long run.

Vigouroux, Vella, and Beis also pointed out the importance of usability testing in disabled groups, demonstrating that multimodal systems combining voice and tactile input by far surpass unimodal systems in terms of user satisfaction and effectiveness [17]. However, their system is still at the level of a conceptual prototype and does not yet provide multimodal integration in embedded settings.

2.3 Research Gap and Justification

Computer vision and deep learning-based gesture recognition is a robust interaction model, best suited for functional upper limb users [11]. Nevertheless, surveys indicate a persistent shortage in the multimodal home automation system design for elderly and disabled users that is comprehensive and embedded. Most existing solutions are single or two-modality based, frequently neglecting gesture-voice, eye-gesture, or full three-way integration, thus not addressing the population with crossing impairments [17], [23]. In addition, employing cloud-based voice APIs, commercial trackers, and expensive sensors limits field deployment in low-resource settings [17], [23]. Therefore, this paper addresses these significant gaps by offering a modular, fully offline, and low-cost multimodal solution developed using open-source libraries and executed on a Raspberry Pi 4 platform. It supports dynamic intramodality switching, local face recognition-based access control, and is verified by deployment with real users in a caregiving center. In so doing, it constitutes a significant step toward the democratization of access to smart space based on inclusive design principles that make assistive technology both technologically adequate and socially equitable [17].

3 Methodology and Design

Designing and engineering an effective, accessible, and low-cost home automation platform has to be addressed as a multi-disciplinary solution entailing real-time embedded computing, multimodal interaction, and user authentication security [12], [23], [24]. The proposed system allows elderly citizens and disabled individuals to control the appliances in a home through three concurrent input modalities: eye gaze, hand gesture, and voice [3], [23], [25]. The platform is hosted on a Raspberry Pi 4 Model B, a compact, low-power computing board supporting real-time computer vision and voice processing algorithm execution [7], [12]. The architecture of the system is composed of five major function blocks:

- Face Recognition for Authentication [12], [25]
- Eye Tracking Interface [12], [26]
- Hand Gesture Recognition [12], [24], [25]
- Speech Recognition Interface [23], [25]
- Control Logic and Device Actuation [3], [23]

All modules are deployed with Python libraries (e.g., OpenCV, MediaPipe, NumPy, dlib, SpeechRecognition) and run in a local, offline environment to preserve user privacy as well as minimize latency [8].

3.1 User Authentication by Face Recognition

One of the security-relevant functionalities of the system is to provide access to only registered users. A face recognition and detection pipeline was established utilizing the Python library face-recognition, which internally employs the dlib framework and Histogram of Oriented Gradients (HOG) with deep learning embeddings (e.g., facial embeddings) [26].

3.1.1 Face Detection

Facial features are being extracted from a live video feed. The model employs pre-trained HOG-based filters to localize facial landmarks and encode them.

3.1.2 Face Recognition

Encoded features are matched against a stored database of registered users based on cosine similarity. Authorized identities are loaded by the SimpleFaceRec module, and encodings map to user names. Upon successful recognition, the user gains access to the control interface. Integration of light frameworks such as LightFace improves recognition accuracy with low computational power [16].

3.2 Eye Movement Analysis

Eye-tracking can allow users to control devices using directional gaze and blink gestures, which become essential for users with limited limb mobility.

3.2.1 Facial Landmark Detection

This work uses the dlib shape predictor model [12] for detecting 68 facial landmarks in the periocular area.

3.2.2 Blink Detection

The EAR is computed by using the vertical and horizontal distances between eye landmarks [12]. A threshold for the EAR, obtained experimentally during the calibration experiments, is used to indicate when a blink occurs.

3.2.3 Gaze Direction Mapping
Gaze tracking is achieved by calculating the pupil centroid within the detected eye region and projecting it relative to the eye boundary. A gaze_ratio value is used to determine if the gaze is left, right, or center. These directional inputs are translated into specific appliance control commands. Figure 1 represents the tested eye direction corresponding to the text and displayed color coding [12].

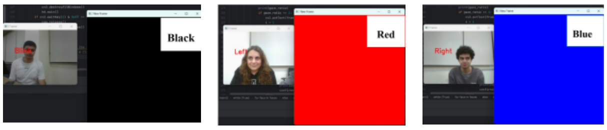

Fig. 1: Simulation results for eye direction correspondences
Source: Generated by the authors

3.3 Hand Gesture Recognition
Gesture-based interaction provides a natural control interface that is specifically adapted to users with unimpaired upper limb mobility. This module was implemented with the innovative real-time hand and finger tracking library named MediaPipe Hands [25].

3.3.1 Hand Landmark Detection
MediaPipe detects 21 3D landmarks per hand without using background subtraction, which greatly improves the accuracy in busy backgrounds.

3.3.2 Finger Counting and Gesture Classification
The position of each raised or lowered finger is detected by reading the vertical positions of key joints, such as landmark 8 vs. landmark 6 for the index finger. Finger counts are translated into discrete appliance commands as shown in Figure 2; for example, 1 finger could turn a light on/off, 2 fingers could operate the curtains, and 5 fingers could operate a fan [25].

Fig. 2: Fingers count display
Source: Created by the authors

3.3.3 Gesture-Based Command Execution
If a gesture is correctly detected, it invokes the associated command. Detection is further assured by visual feedback and logs of the execution of commands, improving user confidence and feedback. It finally maps to the GPIO-controlled outputs: GPIO 17 for lighting, GPIO 27 for fan actuation, and GPIO 18 for curtain control.

3.4 Voice Command Interface
Voice input by visually impaired and/or non-hand-held individuals is mandatory. The system uses Vosk, an offline speech recognition toolkit that can translate spoken commands into executable actions without cloud connectivity.

3.4.1 Audio Capture and Transcription
A USB microphone captures the audio. The audio is processed by a local Vosk model for near-real-time transcription. "Recognizing" and "You said:" kinds of feedback are provided so as to check the responsiveness of the system.

3.4.2 Command Parsing
Recognized phrases are matched with a predefined set of commands, such as "turn on fan" and "open curtain." The system also interprets natural language triggers for modality switching, for example, "go to hand." If speech is not recognized correctly, a message indicates that it is not understandable, as configured in Figure 3.

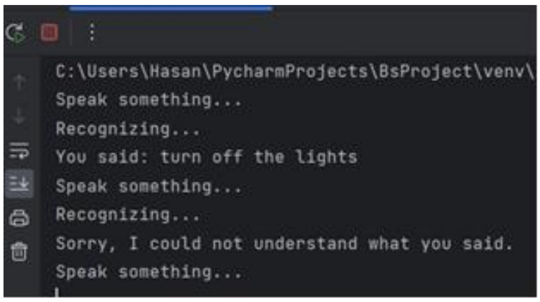

Fig. 3: Voice detection display
Source: Created by the authors

3.5 System Integration and Control Flow
3.5.1 User Authentication
The system starts with face detection and recognition to authenticate the user. Once authenticated, the user is allowed to choose the available control modes: eye movement, hand gestures, or voice commands.

3.5.2 Input Processing
Depending on the mode selected, the system controls the respective inputs, such as eye movements, hand gestures, or voice commands, and then processes them in real time. This real-time processing is necessary for the control of home appliances [24].

3.5.3 Appliance Control

User commands are forwarded to home appliances in this system through relay-based switching devices, which implement the actual physical control required for turning appliances on or off. To make the entire process transparent and comprehensible to the users, every action is assigned a real-time feedback loop. This provides the user with an immediate visual acknowledgment of the status of the system and whether the command has indeed been executed or not. Figure 4 shows the step-by-step control flow, starting from system initialization down to the execution of commands. Initialization includes the users registering with the system, with all their credentials and settings kept locally. They can then choose how they want to control the system: either voice, hand gestures, or eye movement. Each input type is handled in the multimodal framework by its dedicated recognition algorithm. A significant role is played by the ability of the system to dynamically switch between these input modes; a user can override gesture or gaze input at any time, either by pressing the "V" button to switch to voice mode or by using a voice command, "go to hands" or "go to eyes." This flexibility makes the interaction accessible to people with restricted mobility and minimizes physical contact. The way users shut down the system with no risk involves activating the exit control, "X," that returns all appliances to their default states and stops running processes. Therefore, this modular flow allows for context-aware interaction, independence, and reliable fail-safe operation.

Figure 5 depicts the modular architecture employed in implementing the system. The processing-intensive computation tasks of face recognition, gesture detection, and gaze tracking are all handled by a central control of the Raspberry Pi 4 [7], [27]. The long-term data, such as logs, user profiles, and configuration files, are kept on a micro-SD card, making the system adaptable and customizable. The capture of the visual input for authentication or recognition of a gesture or gaze is performed using the Raspberry Pi Camera Module V2 [1], [12]. The USB microphone allows for natural voice interaction, which will be highly beneficial for those users with certain visual or motor impairments [20], [28].

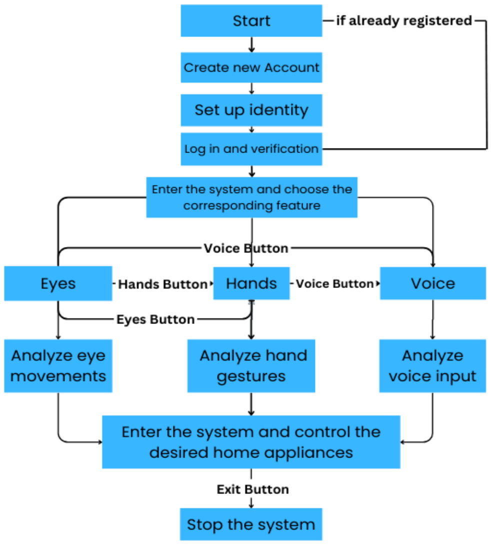

Fig. 4: Control Flow of the system
Source: Created by the authors

The relay modules connect the Raspberry Pi with AC appliances and perform the commands by switching them in real time. This, supplemented by LEDs that indicate the ready state, processing state, and error state, provides additional feedback. The system drives the lamp and fan on GPIO 17 and GPIO 27, respectively, using a 3.3 V-to-5 V level shifter to ensure proper actuation. Employing multimodal input, physical relay control, and layered feedback, this ensures a very user-centric and reliable architecture of the system.

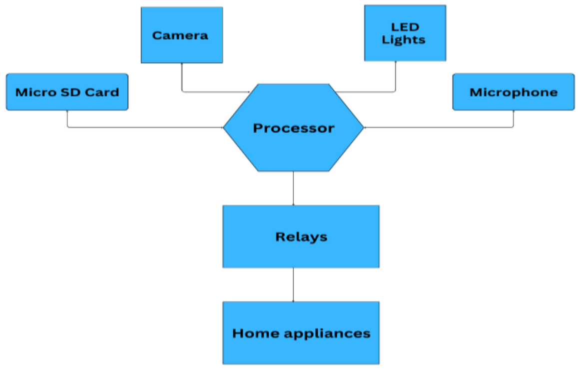

Fig. 5: Blocks for the testing and implementation
Source: Created by the authors

3.6 Hardware Implementation

This system focuses on the integration of the processing unit practically and cohesively with both actuators, sensors, and interface components. Meanwhile, it also seeks to implement some power management techniques to ensure system stability and reduce downtime [29]. The design is modular, responsive in real time, and can handle multiple control modes, making it suitable even for deployment in resource-constrained environments [12], [25]. A key design choice was to utilize modular elements in the form of sensors, relays, and LEDs based on GPIO. This offers ease of upgrade or replacement without necessarily discarding the whole system. Such modularity contributes to long-term sustainability through the reduction of electronic waste while extending the useful life of a device. In that respect, the system addresses a number of Sustainable Development Goals: SDG 3 (Good Health and Well-being), SDG 7 (Affordable and Clean Energy), SDG 10 (Reduced Inequalities), and SDG 12 (Responsible Consumption and Production).

3.6.1 Components of the system and connections

The final implementation of the proposed home automation system integrates multiple hardware components orchestrated through a Raspberry Pi 4 Model B (8 GB RAM, Quad-core ARM Cortex-A72, 1.5 GHz), which serves as the central processing and control unit. As a compact yet powerful single-board computer, the Raspberry Pi is well-suited for real-time execution of parallel tasks, including eye tracking, hand gesture recognition, facial authentication, and voice-based control, all developed using open-source Python libraries [20], [27], [28].

a. Input Interfaces

User input is acquired through multiple peripherals, as represented in Figure 6:

- A Raspberry Pi Camera Module V2 captures facial landmarks for user authentication and gesture patterns for control [12].
- A SunFounder USB Microphone interfaces with the voice command module for voice input acquisition [28].
- Tactile push buttons serve as manual triggers for switching between control modes.
- A MicroSD card provides onboard data storage, enabling lightweight deployment without external servers.

The camera is connected via the Camera Serial Interface (CSI), while the microphone is inserted into a USB port. Initial hardware setup also includes the installation of a cooling fan and heat sink, which ensures thermal stability during prolonged processing. Similarly, the tactile push button permits anticipating any obstacle or failure in automation.

b. Output Interfaces and Actuation:

The system supports three controlled devices:

- Lamp – operated via a Songle relay module.
- Fan – similarly controlled by a dedicated relay module.
- Curtains – actuated using an MG996R servo motor.

To support the 3.3V logic level of the GPIO pins of the Raspberry Pi and the 5V demand of the relay modules, a switching circuit is employed to boost the signal voltage. VCC of the relay is supplied by the 5V rail of the Raspberry Pi, and its control input is driven using this voltage-boosted GPIO link. The appliances are energized to the relay's Common (COM) and Normally Open (NO) terminals, so the appliances will be in an off state by default and will only switch on when explicitly turned on.
For servo actuation, GPIO pin 18 has been used due to its hardware-based PWM support [25]. It provides the precise pulse-width modulation required for the smooth rotation of the MG996R servo motor used for opening and closing the curtain.

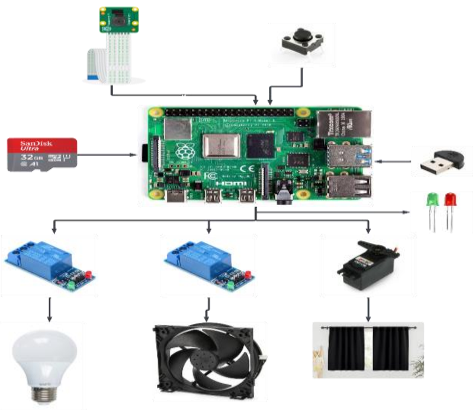

Fig. 6: Final block diagram of the circuitry
Source: Created by the authors

The system status was indicated using two LEDs: a green one for readiness and a yellow one for activity, connected to dedicated GPIO pins. Since the Raspberry Pi 4 uses stable 3.3V logic that fits well with the relay switching circuit, the lamp and fan relays are controlled by GPIO 17 and GPIO 27, respectively. GPIO 18, which supports hardware PWM, was used to drive the MG996R servo motor operating the curtain mechanism. The two status LEDs are connected to GPIO 22 and GPIO 23, which are configured as outputs. These specific assignments help maintain deterministic timing and reduce electrical noise during multimodal operations. Figure 6 gives an overview of the modular and scalable hardware architecture, which, in [13], allows for fully offline functionality while simultaneously processing multiple input modes in real-time. Overall, a powerful and flexible platform for assistive home automation is designed that will have considerable benefits for users with different physical or sensory capabilities.

Takla Sassine, Hassan Mroueh,
Alexandre Yaacoub, Thalia Yammine,
Elias Saadeh, Moustapha El Hassan

3.6.2 Input switching and visual feedback

The three control modalities are integrated at the central system, and the user can dynamically toggle between them by physical push buttons or voice commands [23], represented in Figure 7.

- Button H: Toggles to hand gesture mode
- Button E: Toggles eye tracking
- Button V: Toggles to voice recognition
- Button X: Closes the session
- Real-time status is also reflected by:
- Green LED: Ready to receive commands
- Yellow LED: Command received and in process.

Fig. 7: Connection of the input switching and visual feedback to the central system
Source: Created by the authors

Each push button was assigned to a dedicated GPIO input pin: Hand-gesture toggle (Button H) on GPIO 5, eye-tracking toggle (Button E) on GPIO 6, voice-mode toggle (Button V) on GPIO 13, and the system termination button (Button X) on GPIO 19. All buttons were connected using internal pull-up resistors to prevent undefined states. The LED indicators described above were driven by GPIO 22 (green) and GPIO 23 (yellow), which allow for immediate visual feedback on system readiness and activity. Hence, the control structure provides intermodal switching, and the users can supply input modes concerning context or functional convenience [21].

3.6.3 Final prototype

The following sections describe how the final prototype was built, as represented in Figure 8. All circuit components were first connected and tested. Then, parts for the small house model were 3D-printed, such as supports for the Raspberry Pi Camera Module, windows, doors, and other essential parts shown below [18].

Fig. 8: Final prototype with electrical connections
Source: Created by the authors

The cross-modal assistive system proposed in the present study was benchmarked against a representative cohort of established home-automation frameworks, thereby illustrating its likely impact against the prevailing methods. While extant paradigms have strengthened individual channel performance, pervasive deficits persist in integrated coherence, offline operability, and field-ready execution. Mathew, Sreeshma, Jaison, Pradeep, and Jabarani reported an eye–tracking apparatus which achieved 89% blink detection and 85% gaze fidelity at a nominal 650 ms latency; however, the system omitted user identification, lacked cross-modal validation, and was bound to laboratory confines [12]. Klaib, Alsrehin, Melhem, and Bashtawi expanded modality diversification by integrating gaze data with cloud-based speech interpretation, returning accuracies of approximately 90% and 92%, respectively [23]. Reliance upon proprietary components, namely Tobii sensors and Amazon's Alexa, confined data sovereignty and offline performance, while latency statistics were not stated. In parallel, Harshitaa, Hansini, and Punyo achieved 92% correctness employing an offline convolutional neural network approach to gesture analysis, estimating a latency of 480 ms; adherence to a single-cue channel stands as a limitation [24]. Vigouroux, Vella and Beis provided articulation regarding a dual–mode voice–tactile schema, but the investigation remained resolutely conceptual in nature, yielding neither physical hardware nor quantitative validation [17].

In comparison, the proposed architecture integrates three modalities of eye tracking, gesture recognition, and voice commands within a fully embedded and offline-capable platform on top of low-cost, open-source components. The empirical evaluation achieved accuracies of 93% for eye, 96% for gesture, and 90% for voice, with latency performance competitive across modalities: ~580 ms for eye, ~410 ms for gesture, and ~0.5-1 s for voice. This limitation was mitigated by multimodal redundancy, ensuring accessibility despite environmental or modality-specific failures. Importantly, the system was tested in an ecologically valid caregiving setting, which extends from controlled prototypes to deployment under real-world settings. It surpasses the limitations of past systems in terms of its accuracy and flexibility in realistic applications. This paper, therefore, further contributes to the scientific community in defining the new standard for inclusive assistive automation, which is founded on scientific rigor, ethical grounds, and feasibility in actual operation, as demonstrated by [12], [17].

4 Experimental Validation and Comparative Discussion

The proposed in-built home automation system has been designed to facilitate improved independent living of the disabled and the elderly, using multimodal interaction through eye tracking, hand gesture detection, and voice control, as described in [24], [26]. Secure access is guaranteed through face recognition, with a validation of the system's efficiency. Based on the open-source Raspberry Pi 4 system, the system was tested under real-environment functional testing at the Wahat El Farah Association in El Koura, Lebanon.

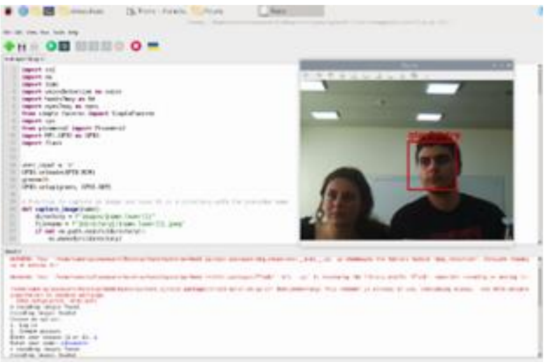

Fig. 9: Face recognition results
Source: Created by the authors

These results confirm that the system can support a broad range of impairments, which is in line with the principles of inclusive design in the context of assistive technology research [17], [19].

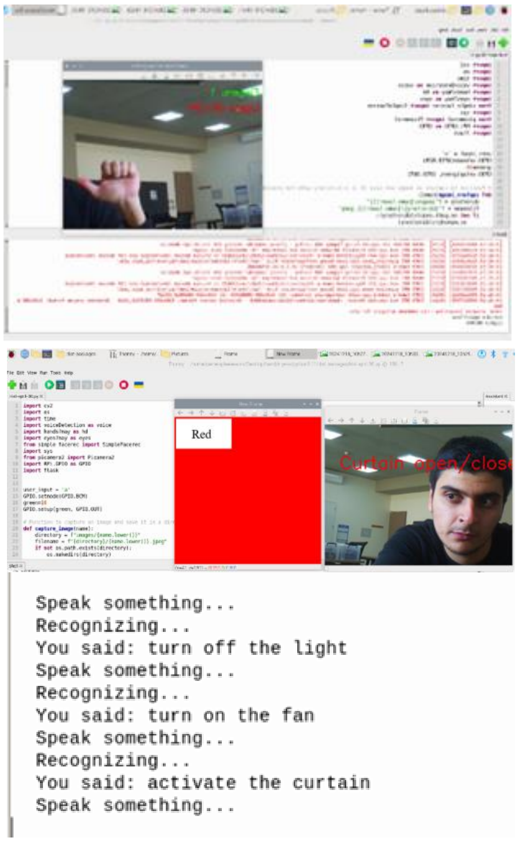

Fig. 10: Gesture recognition, eye tracking, and voice recognition simulation results
Source: Generated by the authors

Referring to Figure 9, the experimental results also show successful face recognition, recognition of hand gestures [24], [25], and effective tracking of eye movements [22], [27]. Finally, voice recognition along with console output is verified [4], [28] as represented in Figure 10.

5 Conclusion

This paper describes a fully embedded, multimodal home automation facility engineered to assist continuity of daily living among elderly users and persons with disabilities. Deployed on Raspberry Pi 4 hardware and leveraged by open-source libraries, the installation interoperates eye tracking, hand gesture interpretation, and vocal directives, enriched by facial biometrics to substantiate secure authentication. A prolonged validated trial in a real caregiving setting exhibited instantaneous rule execution, persistent offline functioning, generalized adaptability, and sustained reliability. The final system effectively addressed physical limitations and substantially raised user independence, dignity, and comfort, outperforming the traditional single-modality assistive technologies. Therefore, the following contributions are determined.

a. Modal Redundancy and Inclusive Co-Interpretation [5], [8].
b. Embedded, Offline-Proxy Architecture [12], [23].
c. Ethical and Economic Integrity [30].
d. Operational Maturity and Deployment Readiness [31], [32].

The proposed multimodal assistive system displays operational feasibility, yet various constraints were uncovered. The vocal interface, implemented using an offline Vosk voice-based model, operates with near real-time latency of approximately 0.5–1 second under controlled conditions. Recognition accuracy decreases under moderate ambient noise levels (55–60 dB), reflecting the trade-off between model compactness and computational constraints on the Raspberry Pi platform [1], [20], [31]. Despite these limitations, the offline implementation ensures privacy, low-latency response, and independence from cloud services, maintaining robust performance in practical home environments. The eye-tracking modality, reliant on pupil-diameter mapping and blink-event identification, demonstrated discerning dependency on head orientation, illuminance, and intrinsic optical alignment; accuracy fell to 81% under luminous conditions characterized by retro-illumination of the pupil. Gesture recognition attained a nominal accuracy of 96.1%, yet remained unsuitable for participants' experiences with profound neuromuscular decline, demonstrated pronounced occlusion risks and displayed performance biases interrogative to demographic factors, including skin pigmentation [19], [33]. Despite execution in residential care settings, the absence of standardized usability indices, such as the System Usability Scale and the NASA Task Load Index, constrained the extrapolability of efficacy conclusions [28], [34].

Consequently, personal responsive care for underrepresented populations, in which future research will be performed, will overcome these limitations through the development of low-power inference accelerators that enable privacy-preserving computation [9], [20], adaptive gaze calibration techniques [35], and the systematic incorporation of biosensors, namely electromyography and inertial measurement units, that collectively broaden system accessibility [12], [19]. Complementary, multimodal sensory and actuator feedback, along with caregiver-facing software applications, will enhance personalization and encourage user self-sufficiency [1], [28], [29], [30].

References:

[1] World Health Organization and World Bank, *World report on disability*, World Health Organization, Geneva, 2011. ISBN: 978-92-4-156418-2.
World report on disability
(Last Accessed December 18, 2025)

[2] Wilson, C., Morris, M. A., and Doughty, K., A scoping review to explore the health, social, and economic outcomes of home automation for people with disabilities, *Journal of Medical Internet Research*, 2023. ISSN: 1438-8871.
National Library of Medicine
(Last Accessed December 18, 2025)

[3] Abedi, A., Yan, H. and Skubic, M., MAISON: Multimodal AI-based sensor platform for older individuals, *Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW)*, 2022, pp. 238–242.
doi.org/10.48550/arXiv.2211.03615

[4] Anggraini, N., Kurniawan, A., Wardhani, L. K., and Hakiem, N., Speech recognition application for the speech impaired using the Android-based Google Cloud Speech API, *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, Vol. 16, No. 6, 2018, pp. 2733–2739. ISSN: 1693-6930.
doi.org/10.12928/TELKOMNIKA.v16i6.9638

[5] HomeCare Magazine, The evolution of smart home technology for people with disabilities, *HomeCare Magazine*, 2024. ISSN: 0899-8205.
(Last Accessed December 18, 2025)

[6] Stella Residence Editorial Team, From ancient doors to smart homes: The complete history of home automation (1900s–2023), *Stella Residence*, 2024.
(Last Accessed December 18, 2025)

[7] Sajay, A, Hong, C, John, W, Determinants of information communication and smart home automation technology adoption for aging-in-place, *Journal of Enabling Technologies*, Vol. 14, Issue 2, 2020, pp.73-86. ISSN: 2398-6263.
doi.org/10.1108/JET-11-2019-0050

[8] Claire, H., Jenny, C., Patricia, W., Kisani, M., Kate, L., Calculating the social impact of home automation for people with disability: A social return on investment study, *Australina Occupational Therapy Journal*,

2024. ISSN: 1440-1630. doi.org/10.1111/1440-1630.12977

[9] Hijazi, S., Lebanon faces the challenge of a rapidly growing elderly population, *L'Orient-Le Jour*, 2021. ISSN: 1935-7258. (Last Accessed December 18, 2025)

[10] United Nations, Department of Economic and Social Affairs, Population Division, *World population ageing 2019: Highlights*, United Nations, New York, 2019. ISBN: 978-92-1-148325-3. (Last Accessed December 18, 2025)

[11] Combaz, E., Situation of persons with disabilities in Lebanon, *Independent Researcher and Consultant in Social Sciences*, 2018, pp. 1–20. Link (Last Accessed December 18, 2025)

[12] Mathew, S., Sreeshma, A., Jaison, T. A., Pradeep, V. and Jabarani, S. S., Eye movement-based cursor control and home automation for disabled people, *Proceedings of the 2019 International Conference on Communication and Electronics Systems (ICCES)*, 2019, pp. 1422–1426. doi:10.1109/ICCES45898.2019.9002325

[13] Hoy, M. B., Alexa, Siri, Cortana, and more: An introduction to voice assistants, *Medical Reference Services Quarterly*, Vol. 37, No. 1, 2018, pp. 81–88. ISSN: 0276-3869, DOI: doi.org/10.1080/02763869.2018.1404391

[14] Yi Jiao, T., Nadine Andrea, F., Félix, P., Delphine, R., Schwab, Tenzin, W., Benefits and barriers associated with the use of smart home health technologies in the care of older persons: A systematic review, *BMC Geriatrics*, Vol. 24, 2024, Article 152. ISSN: 1471-2318, doi.org/10.1186/s12877-024-04702-1

[15] Rocha, A. P., Ketsmur, M., Almeida, N. and Teixeira, A., An accessible smart home based on integrated multimodal interaction, *Sensors*, Vol. 21, No. 16, 2021, Article 5464. ISSN: 1424-8220, doi.org/10.3390/s21165464

[16] Serengil, S. I. and Ozpinar, A., LightFace: A hybrid deep face recognition framework, *Proceedings of the 1st IEEE ASYU Conference*, 2020, pp. 1–5. doi.org/10.1109/ASYU50717.2020.9259802

[17] Vigouroux, N., Vella, F. and Beis, J. M., Usability study of tactile and voice interaction modes by people with disabilities for home automation controls, *arXiv preprint*, arXiv:2211.13042, 2022, pp. 1–12. doi.org/10.48550/arXiv.2211.13042

[18] Dinesh, K., Jaswanth, R., Hareesh, K., Smart Home Automation System, *International Research Journal on Advanced Engineering Hub (IRJAEH)*, 2024, pp. 1122–1128. ISSN: 2584-2137, doi.org/10.47392/IRJAEH.2024.0155

[19] Cleland, J., Wilson, C., Morris, M. A. and Doughty, K., A scoping review to explore the Health and social outcomes of home automation for people with disabilities: A scoping review, *Disability and Rehabilitation: Assistive Technology*, 2023, pp. 1–15. ISSN: 1748-3107, DOI: doi.org/10.1080/17483107.2023.2196308

[20] Wenda, L., Tan, Y., Isil, E., Aaron, L., Motivations, barriers and risks of smart home adoption: From systematic literature review to conceptual framework, *Energy Research & Social Science*, Vol. 80, 2021, Article 102211. ISSN: 2214-6326, doi.org/10.1016/j.erss.2021.102211

[21] Soojung, C., Kyeongsook, N., Smart home adoption: The Impact of user characteristics and perceived benefits, *Buildings*, Vol. 11, No. 9, 2021, Article 393. ISSN: 2075-5309, DOI: doi.org/10.3390/buildings11090393

[22] Babu, V. and Sivasubramanian, S., Home automation through eyeball motion for disabled persons, *International Journal of Engineering and Advanced Technology*, Vol. 8, No. 6S3, 2019, pp. 116–119. ISSN: 2249-8958. doi:10.35940/ijeat.F1406.0986S319

[23] Klaib, A. F., Alsrehin, N. O., Melhem, W. Y. and Bashtawi, H. O., IoT smart home using eye tracking and voice interfaces for the elderly and special needs, *Journal of Communications*, Vol. 14, No. 7, 2019, pp. 614–621. ISSN: 1796-2021, doi:10.12720/jcm.14.7.614-621

[24] Harshitaa, A., Hansini, P., and Punyo, A., Gesture-based home appliance control system for disabled people, *Proceedings of the 2021 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 2021, pp. 1501–1505. doi: 10.1109/ICESC51422.2021.9532973

[25] Sung, G., Sokal, K., Uboweja, E., Bazarevsky, V., Baccash, J., Bazavan, E. G., Chang, C., and Grundmann, M., On-device real-time hand gesture recognition, *arXiv preprint*, arXiv:2111.00038, 2021, pp. 1–8. doi.org/10.48550/arXiv.2111.00038

[26] King, D. E., Dlib-ml: A machine learning toolkit, *Journal of Machine Learning Research*, Vol. 10, 2009, pp. 1755–1758. ISSN: 1532-4435, doi: 10.1145/1577069.1755843

[27] Lifeline Systems, Inc., Lifeline Systems, Inc., *Wikipedia*, 2024. (Last Accessed December 18, 2025)

[28] Iancu, B., Evaluating Google Speech-to-Text API's performance for Romanian e-learning resources, *Informatica Economica*, Vol. 23, No. 1, 2019, pp. 17–25. ISSN: 1453-1305, doi.org/10.12948/issn14531305/23.1.2019.02 (Last Accessed December 18, 2025)

[29] Chrakie, A., El Hassan, M. and Daba, J., A power management system, *WSEAS Transactions on Electronics*, Vol. 16, 2025, pp. 60–72. ISSN: 1109-9445, doi: 10.37394/232017.2025.16.7

[30] Howard, A. G., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., and Adam, H., Searching for MobileNetV3, *arXiv preprint*, arXiv:1905.02244, 2019, pp. 1–9. doi.org/10.48550/arXiv.1905.02244

[31] Banbury, C., Zhou, C., Fedorov, I., Matas, R., Thakker, U., Gope, D., Janapa Reddi, V., Mattson, P., and Whatmough, P., MicroNets: Neural network architectures for deploying TinyML applications on commodity microcontrollers, *arXiv preprint*, arXiv:2010.11267, 2020. doi.org/10.48550/arXiv.2010.11267

[32] Côté-Allard, U., Phinyomark, A. and Gosselin, B., Deep learning for electromyographic hand gesture signal classification using transfer learning, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 27, No. 4, 2019, pp. 760–771. ISSN: 1534-4320, DOI: 10.1109/TNSRE.2019.2896269 doi.org/10.1109/TNSRE.2019.2896269

[33] Raphael, A., Ebenezer, A., Assistive technology and smart environments: A review of current trends and future directions in reverse logistics research, *Journal of Thermal Spray Technology* Vol. 36, 2024, pp. 379-408. ISSN: 1059-9630, doi.org/10.1007/s10696-023-09493-y

[34] Zajicek, M., Successful and available: Interface design exemplars for older users, *Interacting with Computers*, Vol. 16, No. 3, 2004, pp. 411–430. ISSN: 0953-5438, doi.org/10.1016/j.intcom.2004.04.003

[35] Zhang, Y., Sugano, Y. and Bulling, A., Evaluation of appearance-based methods and implications for gaze-based applications, *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems,* paper No. 416, 2019, pp. 1–13. ISBN:978-1-4503-5970-2, doi.org/10.1145/3290605.3300646

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

Moustapha El Hassan, Takla Sassine, Hassan Mroueh, Alexandre Yaacoub, and Thalia Yammine conceptualized the research.
Takla Sassine, Hassan Mroueh, Alexandre Yaacoub, and Thalia Yammine developed the methodology.
Moustapha El Hassan, Takla Sassine, Hassan Mroueh, Alexandre Yaacoub, and Thalia Yammine formulated and analyzed all mathematical models.
Hassan Mroueh and Thalia Yammine simulated the proposed algorithm.
Moustapha El Hassan carried out the validation.
Moustapha El Hassan, Takla Sassine, Hassan Mroueh, Alexandre Yaacoub, Thalia Yammine, and Elias Saadeh conducted the formal analysis.
Moustapha El Hassan was responsible for the investigation.
Takla Sassine, Hassan Mroueh, Alexandre Yaacoub, and Thalia Yammine provided the necessary resources and managed data maintenance.
Takla Sassine, Hassan Mroueh, Elias Saadeh, and Moustapha El Hassan wrote the initial draft.
Elias Saadeh and Moustapha El Hassan reviewed and edited the manuscript.
Elias Saadeh and Moustapha El Hassan prepared the visualizations.
Moustapha El Hassan was responsible for supervision.
Moustapha El Hassan managed the project.

All authors have read and agreed to the published version of the manuscript.

Sources of Funding for Research Presented in a Scientific Article or the Scientific Article Itself

The University of Balamand funded this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.