Vision-Controlled Hand Gesture Recognition System for Home Automation

Luis Gilberto Rangel Pérez
Centro Universitario de los Valles
Universidad de Guadalajara
Ameca, Jalisco, México
Email: luis.rangel4328@alumnos.udg.mx

Rodolfo Omar Domínguez García
Centro Universitario de los Valles
Universidad de Guadalajara
Ameca, Jalisco, México
Email: odomi@academicos.udg.mx

Miriam González Dueñas
Centro Universitario de los Valles
Universidad de Guadalajara
Ameca, Jalisco, México
Email: miriam.gduenas@academicos.udg.mx

Miguel Angel de la Torre Gómora
Centro Universitario de los Valles
Universidad de Guadalajara
Ameca, Jalisco, México
Email: miguel.dgomora@academicos.udg.mx

***Abstract*— The increasing need to support elderly individuals and people with disabilities has driven the development of innovative assistive technologies that enhance independence and improve quality of life. This manuscript presents a gesture-controlled home automation system that allows users to intuitively operate household devices, such as lights and blinds, without physical contact. The architecture of the proposed system is supported by a Raspberry_Pi 4, with OpenCV and MediaPipe for real-time hand gesture recognition. The system interprets commands and executes corresponding actions on connected devices, taking advantage of finger positions. Unlike traditional automation solutions that rely on voice commands or touch interfaces, this approach offers an accessible and hygienic alternative, particularly beneficial for individuals with mobility or speech impairments. The system exhibits high accuracy, rapid response times, and user-friendly operation, making it a cost-effective and scalable solution for smart home applications. This is achieved through the integration of low-cost devices and open-source libraries. Through this work, we contribute to the advancement of inclusive and contactless technology, fostering greater autonomy and accessibility in everyday living environments.**

Keywords- Assistive technology, hand gesture recognition, home automation, Raspberry_Pi, accessibility.

I. Introduction

Jesse Jackson, a civil rights activist, said that "inclusion is the key to growth". For this reason, Assistive Technology (AT) is being increasingly promoted worldwide, as it encompasses equipment, devices, tools, and software designed to reduce barriers, enhance accessibility, and improve access to resources. These technologies empower individuals with hearing impairments, visual impairments, physical disabilities, or progressive loss of autonomy, fostering greater independence, self-determination, and inclusion in daily life [1]. According to the World Health Organization (WHO), there are 1 billion people in the world who need an assistive device, yet only 1 in 10 has access to one. In low- and middle-income countries, only 5% to 15% have access to AT. In Mexico, the 2020 National Census of Population and Housing shows that 20.8 million people - 16.5% of the total population - have some form of disability [2].

To enhance quality of life, the WHO, together with the Global Cooperation in Assistive Technology (GATE) initiative, recognizes access to assistive technologies as a universal right. These technologies play a crucial role in enabling individuals to lead healthy, productive, independent, and dignified lives, facilitating their participation in education, the workforce, and social activities.

AT, serving individuals with disabilities, has significantly contributed to their health and well-being, as well as that of their families. It not only helps prevent the deterioration of their condition but also plays a crucial role in reducing public health expenditures [2]. AT enables and promotes inclusion and participation, especially for people with disabilities, older people and people with non-communicable diseases. The main purpose of these products is to maintain or enhance people's function and autonomy, thus promoting their well-being. They enable people to live a dignified, healthy, productive and independent life, and to learn, work and participate in social life [3].

Currently, one billion people require an assistive device, and this number is projected to exceed two billion by 2030, highlighting the growing global demand of AT. Although anyone can need an assistive product at some point in their lives, the most common users are adults and children with disabilities, older people and people with chronic conditions, such as diabetes and dementia. Examples of assistive or supportive products include hearing aids, wheelchairs, eyeglasses, prostheses and memory aids, among many others. In addition to promoting autonomy and well-being, these products can help prevent or reduce the impact of secondary conditions, such as lower limb amputation in diabetics. They can also reduce the need for and impact of dependency on caregivers and medical and support services. In addition, access to appropriate assistive devices can have major benefits for community development and economic growth [4].

Despite the global need for and recognized benefits of assistive or supportive products, access to them remains limited. Addressing this unmet need is essential to progress towards the Sustainable Development Goals and the implementation of the Convention on the Rights of Persons with Disabilities [5][6].

It also responds to the growing demand for home automation solutions that are accessible and adaptable to the needs of different users. Unlike traditional home automation systems that require touch interfaces or voice commands, gesture-based control offers an innovative alternative that is ideal for people who have difficulty speaking or using physical devices. The COVID-19 pandemic highlighted the importance of non-contact technologies to reduce the spread of disease, and these types of solutions can help improve hygiene and safety in different contexts.

Finally, this project has significant educational and technological value as it combines the use of open-source hardware and software, promoting learning in areas such as computer vision, image processing and embedded control. Its modular and scalable design allows for future improvements and adaptations, making it a basis for the development of more complex automation and accessibility systems.

The project was motivated by the need to help the elderly or people with reduced mobility, who often find it difficult to perform everyday tasks such as turning lights on and off or manipulating objects at different heights. With the advancement of technology, it is possible to develop innovative solutions that improve their quality of life, providing greater autonomy and comfort at home [7].

The present manuscript presents the design and implementation of a control system based on hand gestures. The system enables intuitive and contactless control of light bulbs and the opening/closing of blinds, enhancing accessibility and ease of use. A Raspberry_Pi 4 is used as a central processing unit, together with the OpenCV [8] and MediaPipe libraries [9] for real-time gesture recognition and analysis. Through finger position recognition, the system interprets various commands and performs specific actions on the connected devices, providing an efficient and accessible user experience. The Raspberry_Pi 4 was chosen for its processing capacities, low power consumption, and versatility in integrating with various sensors and actuators. In addition, its compatibility with Python [10] and specialized computer vision libraries, such as OpenCV and MediaPipe, allows the development of optimized algorithms for real-time image detection and processing. These elements ensure accurate and reliable operation, even in environments with lighting variations or unconventional hand positions.

In addition to its application in the home, this project has the potential to be extended to other areas such as industrial automation, advanced domotics, and accessibility systems for people with disabilities. The ability to control devices through gestures without the need for physical contact represents a safe and hygienic alternative, especially in environments where interaction with surfaces must be minimized, such as hospitals or laboratories.

In conclusion, this system aims to provide an innovative, accessible and efficient solution for home automation, improving the independence and quality of life of users. With advanced computer vision and embedded processing technologies, new possibilities are explored in the field of contactless control, with potential applications in various areas of daily life.

The rest of this paper is organized as follows. Section II describes the related work. Section III describes the materials and methodology used in the implementation of the proposed system. Section IV presents and discusses the results, and Section V summarizes the conclusions and future works.

II. Related Work

The interest of the research community in AT is evidenced in a few recent works. For instance, in [11], the current populace of the elderly is apparently abandoned by the younger generations due to their individual circumstances. To enhance vitality and improve the well-being of elderly individuals, an assisted home care system can serve as a valuable solution by offering comprehensive nursing care and continuous monitoring. The proposed system used voice and gesture (MPU6050 accelerometer) to control home appliances like turning on/off the light, closing/opening of curtains, TV, and fan or AC within the living spaces. The system also monitors real-time activities like heart rate and body temperature for the elderly citizens. In the event of an emergency, such as anomalous behaviors indicating a stroke, the proposed system automatically triggers an alarm and turns on an emergency light to alert family members or caregivers. This smart environment can set the temperature and help control the living parameters based on the users' comfort and their health conditions [11]. Gourob et al. [12] reveal that Human-Robot Interaction (HRI) has become an important topic in today's robotic world, especially in assistive robotics. Vision based hand recognition systems provide solutions for these types of human demands. In the system proposed in [12], users do not need to physically operate any devices. Instead, a camera captures hand movements, and these recordings serve as input for gesture-based control, enabling seamless and contactless interaction with the system. Regarding the elderly people, patients and disabled people can benefit from this kind of gesture control. Besides, AT can be used in extreme environments where direct contact is impossible or impractical. Here, a vision-based hand gesture system for controlling robotic hands has been developed. The main intention in [12] is to implement a vision-based gesture recognition system that recognizes data gathered from hand movements through a camera. In [13], the authors reviewed the sign language research in the vision-based hand gesture recognition system from 2014 to 2020. They identified progress and relevant needs that require more attention. The review shows that vision-based hand gesture recognition research is an active field of research, with many studies conducted, resulting in dozens of articles published annually in journals and conference proceedings. Most of the articles focus on three critical aspects of the vision-based hand gesture recognition system, namely: data acquisition, data environment, and hand gesture representation. They also reviewed the performance of the vision-based hand gesture recognition system in terms of recognition accuracy. For the signer dependent, the recognition accuracy ranges from 69% to 98%, with an average of 88.8% among the selected studies. The lack in the progress of continuous gesture recognition could indicate that more work is needed towards a practical vision-based gesture recognition system [13].

The authors in [14] propose a Home Automation System with hand gesture recognition based on Mediapipe, using Arduino UNO. As one gets older, his/her mobility tends to decrease. Therefore, simple tasks such as getting up to switch the lights on or turning the fan off can become difficult. Thus, it became imperative to create a system which allows them to perform these tasks - a “Hand Recognition based Home Automation System”. Various methods for gesture recognition have been discussed as well as how every model produces a varying training time and accuracy. Although an average accuracy was found to be 98% with the majority of the Machine Learning (ML) models and MediaPipe's technology, our suggested methodology demonstrates that MediaPipe with Dense may be effectively utilized as a tool to correctly recognize complicated hand gestures. Additionally, training this particular model takes much less time compared to the other models. Faster real-time detection of gestures demonstrates the efficiency of the model. In [15], the authors propose a remote-control method based on one-handed gestures for mobile manipulator, so that the operator can control the entire robotic system with only one hand. In this study, they combined real-time hand key points detection technology provided by MediaPipe, with the RealSense D435i depth camera to address the inaccuracy in depth recognition problem of the original method. Then, the position, pitch, and rotation of the hand are analyzed to generate control commands. A lightweight gesture recognition model based on a Gated Recurrent Unit (GRU) is proposed to use specific gestures for switching between controlled objects.

In [16], the authors focus on the design of a gesture-controlled robotic arm that is navigated with the help of a webcam and OpenCV-enabled real-time hand tracking. Embedded with OpenCV’s hand tracking module, the system successfully identifies the hand landmarks from the live camera captured hand movement of the user. The state of each finger is represented as a string that includes the $ symbol and a number, such as $00010. This is transmitted to an Arduino UNO board, and, based on this information, the

robotic arm can move its fingers up or down. Control of the flows of the movement of the robotic arm is done using the Arduino UNO whereby the servo motors receive signals to move 0 or 180 degrees. This is mainly due to the efficiency, precision, and short latency of the suggested system, which makes it novel. It gives the user a simple graphics interface to control robotic arms without the necessity to know programming or have specific hardware. This technology provides an effective solution in robotic manipulation and presents the enormous potential to enhance HRI in different application areas [16].

III. MATERIALS AND METODS

Gesture recognition is a technique that interprets hand or body movements to interact with electronic devices without the need for physical contact. It has a wide range of applications, from video games and augmented reality to home automation and accessibility for people with disabilities.

A. Raspberry_Pi 4 and its features

The Raspberry_Pi 4 [17] is a low-cost, high-performance microcomputer ideal for image processing and embedded control applications. Its key features include:

- Quad-Core ARM Cortex-A72 processor.
- Up to 8 GB of RAM.
- General Purpose Input Output (GPIO) ports for connecting to sensors and actuators.
- Support for Python and computer vision libraries.

Raspberry_Pi 4 is used to:

1. Capture real-time video from a Pi or USB camera.
2. Process the images using OpenCV and MediaPipe to recognized gestures.
3. Send signals via the GPIO ports to the actuators that control, for example, spotlights and blinds.

In this system, GPIO ports are used to switch spotlights on and off using relays or voltage control modules and control a motor to open and close blinds based on detected gestures. Communication between the software and the actuators is achieved by programming the Raspberry_Pi 4 in Python, allowing for easy and efficient integration. This can be seen in Figures 1 and 2.

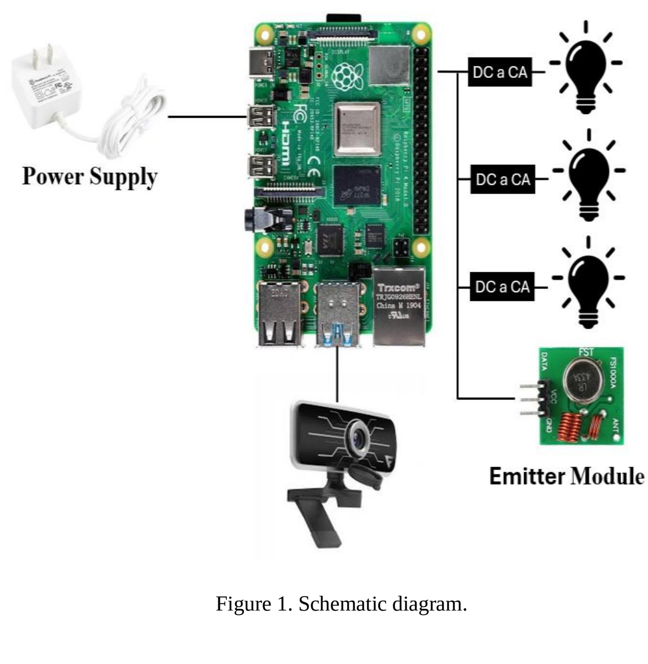

Figure 1. Schematic diagram.

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-268-5

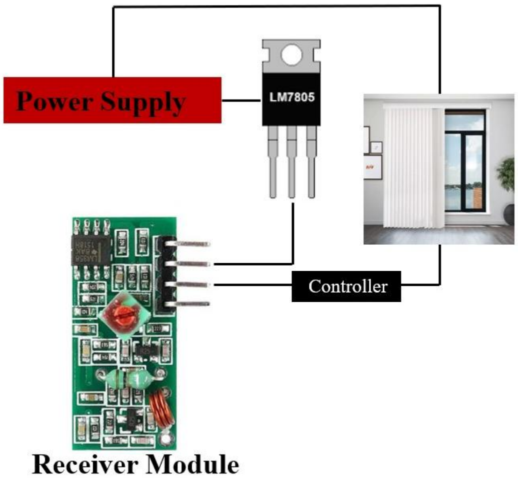

Figure 2. Module actuator blinds.

Figures 3 and 4 show the prototype implemented to carry out the tests and observe how the programmed gestures interact with the system. With the corresponding gesture, the lights are turned on or off accordingly. With other gestures, the blinds are closed or opened (the motor turns on or off).

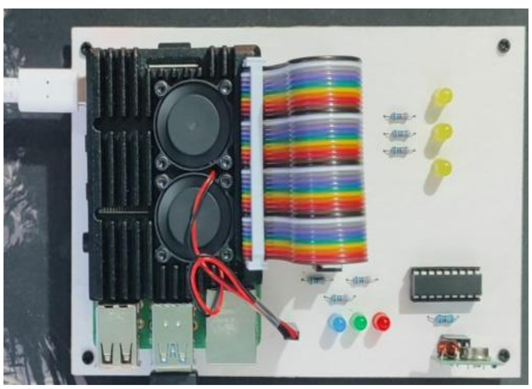

Figure 3. Prototype based on Raspberry_Pi 4B+.

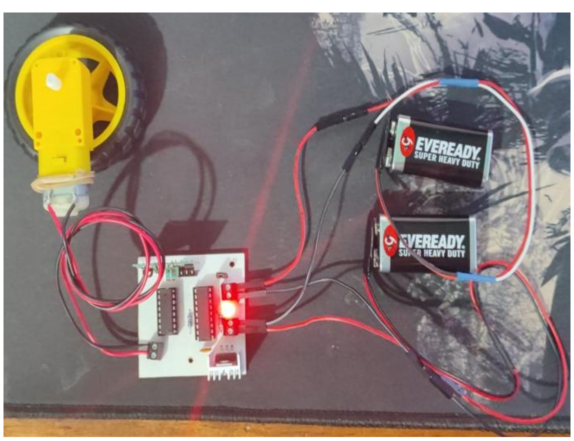

Figure 4. Prototype actuator for blinds.

B. *Gesture recognition techniques*

There are several techniques for recognizing and interpreting gestures, of which the following stand out:

- Sensor-based: These use devices such as accelerometers, gyroscopes and infrared sensors to detect movement.
- Computer vision-based: They use cameras and image processing algorithms to identify the position and movement of the hand.
- Hybrid: Combining physical sensors with computer vision for greater accuracy.
- This project: uses a technique based on computer vision because it enables gesture recognition without the need for additional physical devices.
- Two key libraries are used to implement gesture recognition in this project:
  OpenCV (Open-Source Computer Vision Library): It provides tools for real-time image acquisition, processing and analysis.
  MediaPipe: A framework developed by Google that uses artificial intelligence and machine learning to efficiently recognize hands, faces and poses. MediaPipe provides pre-trained hand recognition models that enable segmentation and analysis of finger movements with high accuracy.

The ability to perceive the shape and motion of hands can be a vital component in improving the user experience across a variety of technological domains and platforms. While coming naturally to people, robust real-time hand perception is a decidedly challenging computer vision task, as hands often occlude themselves or each other (e.g., finger/palm occlusions and handshakes) and lack high contrast patterns [9].

MediaPipe Hands is a high-fidelity hand and finger tracking solution. It employs Machine Learning (ML) to infer 21 3D landmarks of a hand from just a single frame. Whereas current state-of-the-art approaches rely primarily on powerful desktop environments for inference, our method achieves real-time performance on a mobile phone, and even scales to multiple hands [9].

After the palm detection over the whole image, our subsequent hand landmark model performs precise keypoint localization of 21 3D hand-knuckle coordinates inside the detected hand regions via regression, that is direct coordinate prediction, as shown in Figure 5.

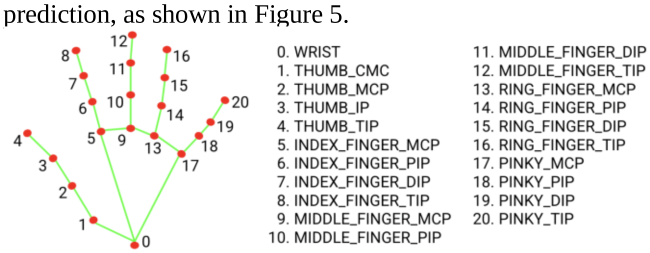

Figure 5. Hand landmarks (taken from [9]).

Figure 6 shows examples of hand gestures that MediaPipe and its training model are able to detect and generate [9].

Figure 6. Examples of hand gestures (taken from [9]).

According to Table I, gesture A is the first action to set the *system enable* state to recognize the on and off orders of the devices. This leads to gestures B or C, which indicate that a device wants to be turned on or off. The system then waits for any gesture from D to H to turn on or off the chosen device. Once the system receives gesture from D to H, and the required action is given, the system returns to the *system enable* state, waiting for gesture A in order to repeat the process.

TABLE I. MENU OF OPTIONS

Gesture	Thumb	Index	Middle	Ring	Pinky	Action
A	0	0	0	0	0	System enable state
B	1	1	1	1	1	It enters you into the power on menu
C	1	0	0	0	1	It takes you to the shutdown menu
D	1	0	0	0	0	Spotlight on/off 1
E	1	1	0	0	0	Spotlight on/off 2
F	1	1	0	0	1	Spotlight on/off 3
G	0	1	0	0	1	Motor (blinds)

IV. RESULTS

The prototype was tested in a controlled environment to observe its operation and estimate its accuracy, response time and ease of use. Figure 7 shows the system enable state (gesture A).

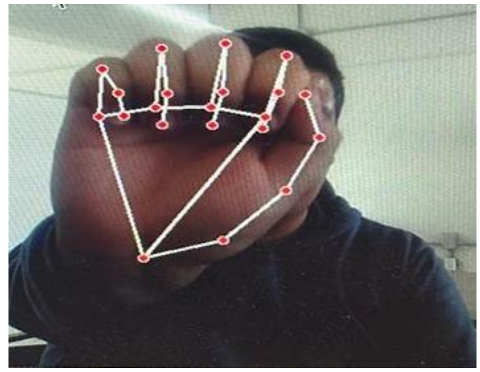

Figure 7. Gesture A, *system enable state.*

Figure 8 shows gesture B, which initiates the power-on menu.

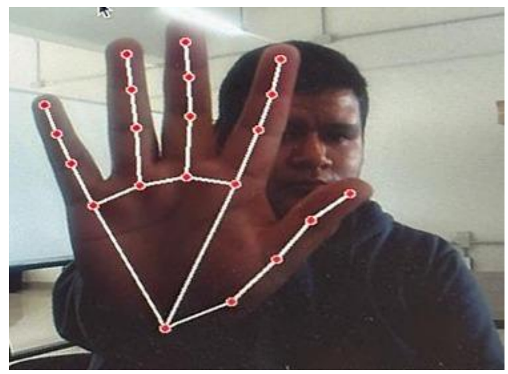

Figure 8. Gesture B grants access to the “power-on menu”.

Figure 9 shows gesture C, which initiates the shutdown menu.

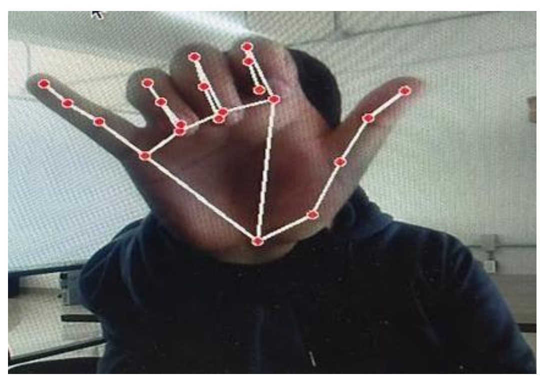

Figure 9. Gesture C, it takes you to the shutdown menu.

Figure 10 shows the gesture D, "Spotlight 1 on/off".

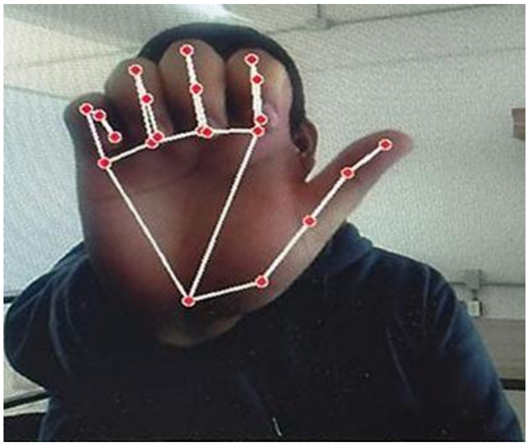

Figure 10. Gesture D, Spotlight 1 on/off.

Figure 11 shows gesture E, "Spotlight 2 on/off".

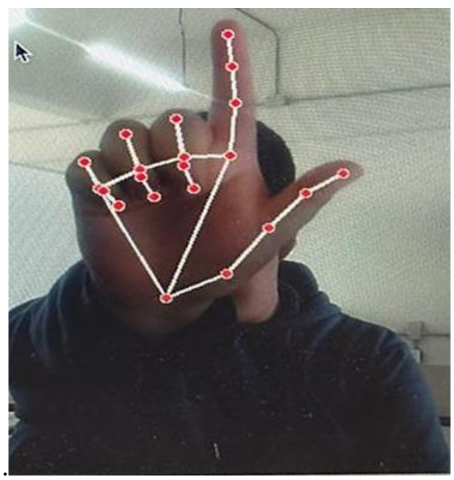

Figure 11. Gesture E, Spotlight 2 on/off.

Figure 12 shows gesture F, "Spotlight 3 on/off".

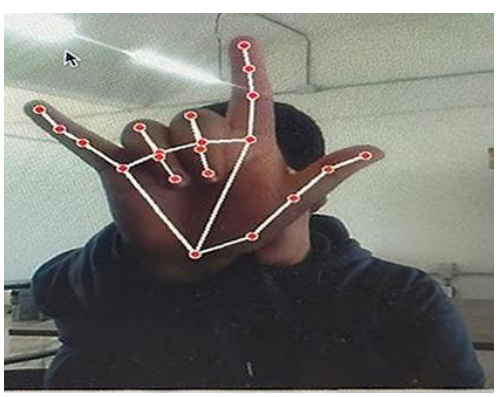

Figure 12. Gesture F, Spotlight 3 on/off.

Figure 13 shows gesture G, "Motor of blind on/off".

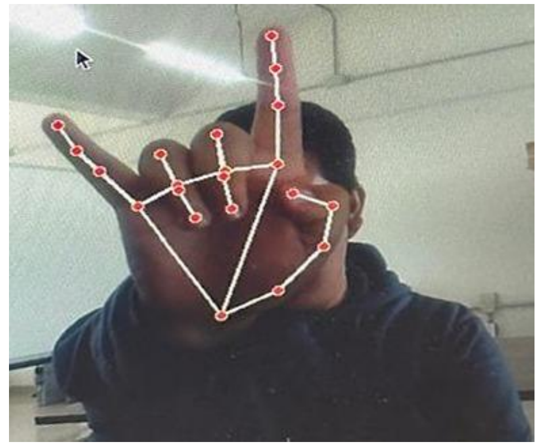

Figure 13. Gesture G, Motor of blind on/off.

Figures 14, 15, and 16 show the process of turning on light bulb 1. The sequence of gestures is a) set the system to enable state, b) enter the system of devices to turn on or off and c) turn on bulb 1.

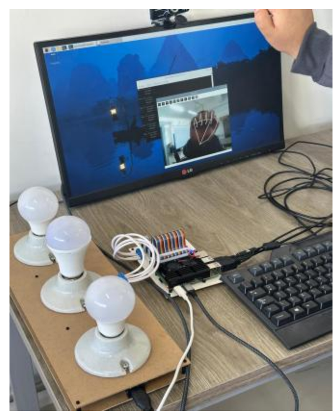

Figure 14. Gesture A, powering on the system.

Figure 15. Gesture B grants access to the "power-on menu".

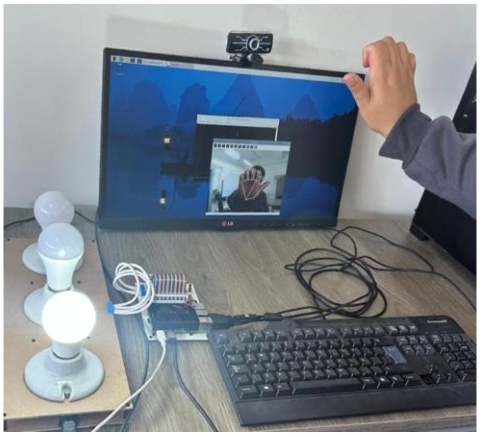

Figure 16. Gesture D, Spotlight on/off.

A. Aspects evaluated

Gesture recognition accuracy: It was verified that gestures were correctly recognised under different lighting conditions and viewing angles.

Response time: The time taken by the system to perform an action after detecting a gesture is almost instantaneous. The response time was fast enough to provide a smooth user experience.

Usability: The user's experience of interacting with the system was evaluated to ensure it was intuitive and did not require technical skills. It was tested with two elderly people, so further testing is needed.

Test results: A high accuracy in gesture recognition was observed during the evaluation process, being robust under varying lighting conditions. On the other hand, users found the system easy to use and they appreciated the ability to control devices without physical contact.

B. Difference from other projects

Focus on hand gestures: Unlike other home automation systems that rely on touch interfaces or voice commands, this project focuses on hand gesture recognition, providing a more intuitive and accessible alternative for people who have difficulty using physical devices or speaking.

Low-cost technologies: The use of computer vision technologies (OpenCV and MediaPipe) and a Raspberry_Pi 4 allows for a low-cost and highly efficient implementation, which is not common in other similar systems that typically require more expensive or complex hardware.

Optimized for different conditions: The project focuses on optimizing image processing for different lighting conditions and viewing angles, making it more robust and adaptable compared to other systems that may not perform well in variable environments.

Ease of implementation: The modularity of the system allows for easy implementation and the possibility of adding more devices in the future.

V. Conclusions And Future Work

The implementation of the vision-controlled hand gesture recognition system demonstrated a practical use case of the Raspberry_Pi 4, combined with OpenCV and MediaPipe, to successfully develop an effective contactless control interface. The proposal shows how embedded systems can be applied to detect hand gestures in real time, and to control devices efficiently and accurately. The Raspberry_Pi 4, based on a single-chip System-on-Chip (SoC), enables features, such as processing power, the ability to run a full operating system, excellent support for Python and its libraries, and various connectivity options, making it an excellent choice for more complex and flexible embedded systems.

The use of OpenCV enables efficient image processing, while MediaPipe facilitates the detection and tracking of hand gestures, optimizing interaction with the system. This approach not only highlights the benefits of the Raspberry_Pi 4 as a computing platform, but also opens the door to future extensions in areas such as automation, accessibility and contactless control in a variety of environments, whether domestic, industrial or commercial.

The affordability of the devices used, coupled with the utilization of open-source software libraries, enables the implementation of this device as a Technology Readiness Level 6 (TR6) prototype. This approach facilitates thorough evaluation and supports the development of a commercial prototype at Technology Readiness Level 9 (TR9).

Additional feedback on system usage from the target population will be collected at a later stage to facilitate a comparative analysis with existing market solutions.

Acknowledgment

We thank the Centro Universitario de los Valles (CUValles) of the Universidad de Guadalajara for the space provided in its laboratories. We would also like to thank everyone who supported us to complete this research work, directly or indirectly.

References

[1] Government of Mexico, Assistive Technologies, [Online] Available: https://www.gob.mx/profeco/es/articulos/tecnologias-asistenciales?idiom=es). [retrieved Nov. 12, 2024].

[2] V. A. Carrizal Pérez, Consumer Magazine, Digital Environment, Assistive Technologies. Jul. 21, 2021, vol. 533, ISSN 0185-8874.

[3] Global Accessibility Reporting Initiative (GARI), Learn about Mobile Accessibility: Vision. [Online]. Available: https://www.gari.info/vision.cfm?lang=es [retrieved Jul. 14, 2024].

[4] National Institute of Statistics and Geography (INEGI), Press Release No. 24/21. [Online]. Available: https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2021/EstSociodemo/ResultCenso2020_Nal.pdf [retrieved Jul. 14, 2024].

[5] National Institute of Statistics and Geography (INEGI), Population with limitations or disabilities by federal entity and type of activity performed or mental condition according to sex, 2020. [Online]. Available: https://www.inegi.org.mx/temas/discapacidad/#Tabulados [retrieved Jul. 14, 2024].

[6] World Health Organization (WHO), List of priority technical aids. Geneva, Switzerland, [Online], Available: http://apps.who.int/iris/bitstream/handle/10665/207697/WHO

_EMP_PHI_2016.01_spa.pdf;jsessionid=7643DD191B70DE7312F44E299F211411?sequence=1 [retrieved at Jul. 14, 2024].

[7] World Health Organization (WHO), Assistive Technology, [Online]. Available: https://www.who.int/es/news-room/fact-sheets/detail/assistive-technology. [retrieved Jul. 14, 2024].

[8] OpenCV. [Online]. Available: https://opencv.org/ [retrieved Jul. 14, 2024].

[9] MediaPipe Hands. [Online]. Available: https://github.com/google-ai-edge/mediapipe/blob/master/docs/solutions/hands.md. [retrieved Jul. 14, 2024].

[10] Python. [Online]. Available from: https://www.python.org/ [retrieved Jul. 14, 2024].

[11] H. Basanta, Y. -P. Huang and T. -T. Lee, "Assistive design for elderly living ambient using voice and gesture recognition system," 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, 2017, pp. 840-845, doi: 10.1109/SMC.2017.8122714

[12] J. Hossain Gourob, S. Raxit and A. Hasan, "A Robotic Hand: Controlled With Vision Based Hand Gesture Recognition System," 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI), Rajshahi, Bangladesh, 2021, pp. 1-4, doi: 10.1109/ACMI53878.2021.9528192.

[13] N. Mohamed, M. B. Mustafa and N. Jomhari, "A Review of the Hand Gesture Recognition System: Current Progress and Future Directions," in *IEEE Access*, vol. 9, pp. 157422-157436, 2021, doi: 10.1109/ACCESS.2021.3129650.

[14] O. Mishra, P. Suryawanshi, Y. Singh, and S. Deokar. A Mediapipe-Based Hand Gesture Recognition Home Automation System. In 2nd International Conference on Futuristic Technologies (INCOFT), IEEE, pp. 1-6, 2023.

[15] J. Xie et al., "One-Handed Wonders: A Remote Control Method Based on Hand Gesture for Mobile Manipulator," 2024 International Conference on Advanced Robotics and Mechatronics (ICARM), Tokyo, Japan, 2024, pp. 686-691, doi: 10.1109/ICARM62033.2024.10715857.

[16] K. Singh, A. Joshi, J. Bhatt, A. Juyal and M. S. Rawat, "Webcam-Guided Gesture-Controlled Robotic Arm," 2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC), Gwalior, India, 2024, pp. 1297-1303, doi: 10.1109/AIC61668.2024.10730972.

[17] Raspberry Pi. [Online]. Available: https://www.raspberrypi.com/products/raspberry-pi-4-model-b/. [retrieved Jul. 14, 2024].