International Journal for Modern Trends in Science and Technology
Volume 11, Issue 05, pages 97-104
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol11issue05.html
DOI: https://doi.org/10.5281/zenodo.15266766

AI-Based Gesture & Voice – Controlled Home Automation System

K. Srinivasa Rao, G. Namratha Phoebe, K. Sri Satya Sai Naga Durga, J. Anandhitha, K. Madhuri

Department of Electronics and Communication Engineering, Andhra Loyola Institute of Engineering and Technology, Andhra Pradesh, India.

To Cite this Article
K. Srinivasa Rao, G. Namratha Phoebe, K. Sri Satya Sai Naga Durga, J. Anandhitha, K. Madhuri (2025). AI-Based Gesture & Voice – Controlled Home Automation System. International Journal for Modern Trends in Science and Technology, 11(05), 97-104. https://doi.org/10.5281/zenodo.15266766

Article Info
Received: 22 March 2025; Accepted: 18 April 2025.; Published: 20 April 2025.

KEYWORDS

Gesture Control, Offline Voice Control, Home Automation, MediaPipe, Vosk AI, Raspberry Pi, Energy Optimization, Smart Home, Safety Monitoring

ABSTRACT

This paper presents a robust and efficient Gesture-Based and Voice (AI-Offline) Controlled Home Automation System aimed at enhancing convenience, energy optimization, and safety within a smart home environment. The proposed system utilizes a Raspberry Pi 3 as the core processing unit, integrated with various sensors and modules including a DC fan, a bulb, a microphone, a relay module, a camera module, a motion sensor (IR), a gas sensor, and a fire/smoke sensor. Gesture control is achieved through the MediaPipe framework, which captures and processes hand gestures via the camera module to authorize and control appliances. Additionally, offline voice control is implemented using Vosk AI, enabling voice commands to operate the fan and bulb without the need for internet connectivity.

The system incorporates three essential sensors to enhance safety by detecting gas leaks, fire, and motion, which are promptly reported through a Telegram bot with image alerts. Furthermore, energy optimization is achieved by automatically switching off appliances when no motion is detected for a specified duration. The integration of gesture control, offline voice recognition, and sensor-based safety mechanisms demonstrates a user-friendly and reliable approach for smart home automation. The proposed system offers substantial potential for practical applications in energy-efficient, secure, and user-centric smart home environments.

1. INTRODUCTION

In recent years, smart home automation systems have gained significant attention due to their potential to enhance convenience, energy efficiency, and safety. The rapid advancements in the Internet of Things (IoT), Artificial Intelligence (AI), and embedded systems have contributed to the development of innovative solutions for controlling home appliances using voice, gestures, and mobile applications. Traditional home automation systems primarily rely on internet connectivity for seamless operation, limiting their accessibility and robustness in environments with poor network availability. This paper proposes a Gesture-Based and Voice (AI-Offline) Controlled Home Automation System, designed to provide an efficient, reliable, and user-friendly experience for home automation without the dependency on internet connectivity. The system utilizes a Raspberry Pi 3 as the central processing unit, integrated with various sensors and modules to enhance control and safety mechanisms within the smart home environment.

• The proposed system features two primary control methods: Gesture Control and Offline Voice Control. Gesture recognition is implemented using the MediaPipe framework allowing users to operate home appliances through hand gestures captures by a camera module. Additionally, the Vosk AI framework is employed to enable offline voice control, providing a flexible solution for users to control devices using predefined voice commands without requiring an active internet connection.Moreover, the system integrates several sensors, including motion, gas, and fire/smoke sensors, to enhance safety monitoring and ensure timely alerts through a Telegram bot. A unique aspect of this project is the implementation of an energy optimization mechanism, where appliances are automatically switched off when no motion is detected for a specific duration.

The proposed system aims to bridge the gap accessibility and reliability in smart home automation by offering a dual-mode control mechanism. This paper discusses the design, implementation, and performance evaluation of the system, providing insights into its practical applications in modern.

1.1. Objectives:

- To develop a smart home automation system that enables control of household appliances using gesture-based recognition and offline voice commands.
- To integrate AI-driven gesture recognition using Mediapipe technology for efficient and accurate gesture detection.
- To implement Vosk AI for offline voice control, ensuring reliable performance without internet dependency.
- To enhance safety and security through the inclusion of gas, fire/smoke, and motion sensors, providing real-time alerts via Telegram bot.
- To optimize energy consumption by automatically turning off appliances when no motion is detected for a specified duration.
- To create a user-friendly interface and reliable system that operates seamlessly for home automation

1.2.Principles:

- Gesture Recognition using Mediapipe:

• Mediapipe is an open-source framework developed by Google, which provides advanced tools for real-time perception and processing of multimodal data.

• It utilizes machine learning models to detect hand gestures accurately and process them for triggering predefined actions.

• The framework provides high-performance inference and processing through pipelines that handle image and audio inputs, making it suitable for real-time gesture-based control.

- Offline Voice Recognition using Vosk AI:

• Vosk AI is an offline speech recognition toolkit that allows speech-to-text processing without requiring an internet connection.

• It is designed to recognize voice commands efficiently with high accuracy, making it ideal for scenarios where connectivity may be unreliable or unavailable.

• The voice control system responds to predefined commands to turn appliances ON or OFF, ensuring robust control.

- Sensor Integration for Safety & Energy Optimization:

• Various sensors, including Gas Sensors, Fire/Smoke Sensors, and IR Motion Sensors, are integrated to enhance safety and efficiency.

• The Gas and Fire/Smoke Sensors continuously monitor environmental conditions, and if a threat is detected, an alert is sent to a Telegram bot along with a captured image.The IR Motion Sensor is employed to detect occupancy and automatically turn off appliances

1.3 Processes Involved:

The proposed Gesture-Based & Voice (AI-Offline) Based Home Automation System involves several interconnected processes that work together to achieve reliable and efficient control of household appliances. The key processes are:

1.3.1 Gesture Recognition Process:

• The gesture recognition system uses the MediaPipe framework, a powerful tool for real-time hand tracking and gesture recognition.

• The camera module connected to the Raspberry Pi 3 captures live video feeds.

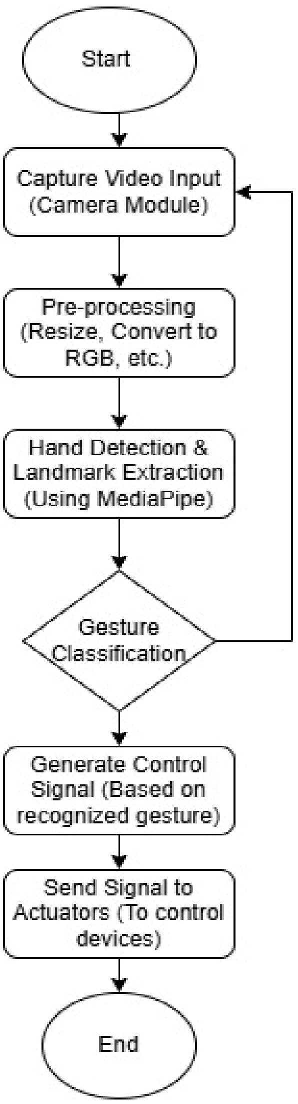
• MediaPipe processes these images to detect hand landmarks and identify specific gestures.

• Detected gestures are compared with predefined gestures (e.g., swiping left, swiping right) to perform corresponding actions such as turning appliances ON or OFF.

• Gesture data is then transmitted to the Raspberry Pi, which triggers the relay module to control the connected devices accordingly.

1.3.2 Offline Voice Recognition Process:

• The offline voice recognition system is implemented using Vosk AI, an open-source speech recognition toolkit that operates without internet dependency.



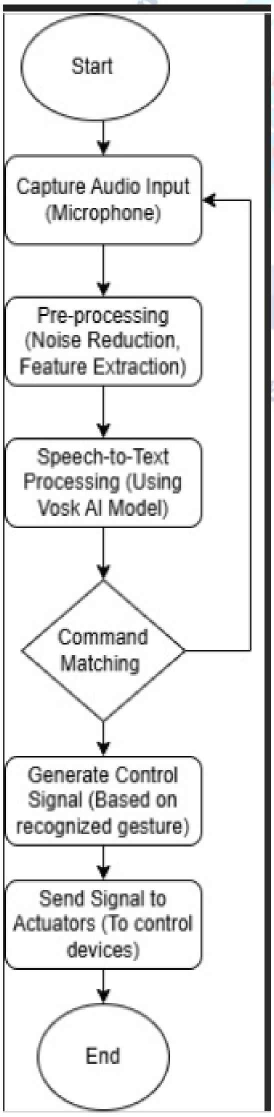


Fig : Gesture control Fig : Voice control

• The microphone connected to the Raspberry Pi 3 captures audio commands.

• Vosk AI processes these commands to recognize specific keywords related to controlling devices (e.g., "Turn on the fan", "Turn off the bulb").

• The recognized commands are translated into control signals and sent to the relay module for execution.This approach ensures privacy, reliability, and accessibility, even when internet connectivity is not available.

1.3.3 Sensor Monitoring and Safety Mechanism:

• Three sensors are integrated to enhance safety monitoring:

· Gas Sensor: Continuously monitors for the presence of hazardous gases.

· Fire/Smoke Sensor: Detects smoke or fire in the surroundings.

· IR Motion Sensor: Detects human motion to determine if the room is occupied.

• If a gas leak or fire is detected, the Raspberry Pi immediately sends an alert to a Telegram bot, including a picture captured by the camera module.

• The IR Motion Sensor is also used for energy optimization by turning off appliances automatically when no motion is detected for a certain duration.

1.3.4 Communication through Telegram Bot:

• A Telegram bot is configured to provide real-time alerts when the system detects potential threats (e.g., gas leak, fire, or intrusions).

• The camera module captures images when an alert is triggered and sends them to the user via the bot.

• This ensures prompt notification and appropriate action from the user even when they are away from home.

1.3.5 Energy Optimization Mechanism:

• The IR Motion Sensor plays a crucial role in optimizing energy usage.

• When no motion is detected for a predefined period, the Raspberry Pi sends signals to the relay module to switch off connected appliances automatically.

• This process helps reduce unnecessary power consumption and promotes energy efficiency.

2. MATERIALS & METHODS

This section outlines the various hardware and software components utilized in developing the Gesture-Based &

Voice (AI-Offline) Based Home Automation System. It also details the integration process of these components to achieve seamless and efficient operation.

2.1 Hardware Components

The hardware components used in this project include various sensors, processing units, and other essential devices to ensure the reliable functioning of the proposed system. The key hardware components are:

• Raspberry Pi 3:

· Acts as the central processing unit for the entire system.

· Facilitates communication between the sensors, gesture recognition module, and voice recognition module.

· Utilizes its GPIO (General Purpose Input/Output) pins to connect and control various devices.

• Camera Module:

· Used for capturing real-time hand gestures.

· Connected to the Raspberry Pi 3 for data processing and recognition using the MediaPipe framework.

· Provides high-resolution image capturing capability for accurate gesture recognition.

• Sensors:

· Motion Sensor (IR Sensor): Detects occupancy and automatically turns off appliances if no motion is detected for a specified duration, thereby promoting energy efficiency.

· Gas Sensor (MQ-2): Detects the presence of harmful gases and alerts the system to trigger necessary safety measures.

· Fire/Smoke Sensor: Monitors the environment for signs of fire or smoke and immediately triggers appropriate actions, such as shutting down appliances or sending alerts.

• Relay Modules:

· Used for switching appliances ON and OFF based on commands from the Raspberry Pi 3.

· Acts as a bridge between the low-power control system (Raspberry Pi) and high-power household appliances.

• Microphones:

· Used to capture voice commands from the user.

· Works in coordination with the Vosk AI framework for processing offline voice commands.

Component	**RaspberryPi(GPIO Pins)**
IR Motion Sensor	GPIO 17 (Pin 11)
Gas Sensor (MQ-2)	GPIO 27(Pin 13) (DO)
Fire/Smoke Sensor	GPIO22(Pin15) (DO)
Relay Module (Fan Control)	GPIO 23 (Pin 16)
Relay Module (Light Control)	GPIO 24 (Pin 18)
Camera Module	CSI Port / USB Port
Microphone	USB port
Power & Ground	3.3/5v and GND pin

Fig : Components Table

2.2 Software Components

The software components consist of various frameworks, programming languages, and communication protocols to enable the system's functionality. The main software tools are:

· Mediapipe Framework:

o An open-source framework developed by Google for real-time perception and processing of multimodal data.

o Utilized for implementing the gesture recognition system.

o Captures live video feeds and processes them to detect hand landmarks and identify specific gestures.

o Compares detected gestures with predefined gestures to trigger corresponding actions such as turning appliances ON or OFF.

· Vosk AI:

o An open-source speech recognition toolkit designed for efficient offline speech processing.

o Converts voice commands into text and compares them with predefined commands to control appliances without requiring internet connectivity.

o Provides a robust solution for voice recognition even in low-resource environments.

· Python Programming Language:

o Used for coding the entire system, including interfacing with sensors, gesture recognition, voice recognition, and communication protocols.

o Facilitates data handling and processing using various Python libraries.

- Telegram Bot API:

o Employed for real-time alerts and notifications.

o Enhances the system's reliability by providing alerts to users even when they are away from home.

- Operating System (Raspberry Pi OS):

o The default operating system used for running the software modules on the Raspberry Pi 3.

o Provides compatibility and seamless integration with various hardware components.

2.3 Integration of Hardware and Software

The integration process involves connecting all the hardware components to the Raspberry Pi 3 and ensuring their compatibility with the software modules. The steps involved are:

- Setting up the Raspberry Pi 3 with the necessary operating system and required libraries for handling camera input, sensors, and voice recognition.
- Configuring the camera module to capture real-time hand gestures and processing them using the Mediapipe framework.
- Implementing voice recognition using Vosk AI to detect predefined commands.
- Linking the sensors (Motion, Gas, and Fire/Smoke) to the Raspberry Pi 3 for continuous monitoring of environmental conditions.
- Establishing communication between the Raspberry Pi 3 and relay modules for controlling household appliances.
- Setting up the Telegram Bot API to send notifications and alerts to the user in case of emergency or required actions.
- Testing the overall system to ensure proper communication between hardware and software components and verifying that the system responds accurately to user inputs.

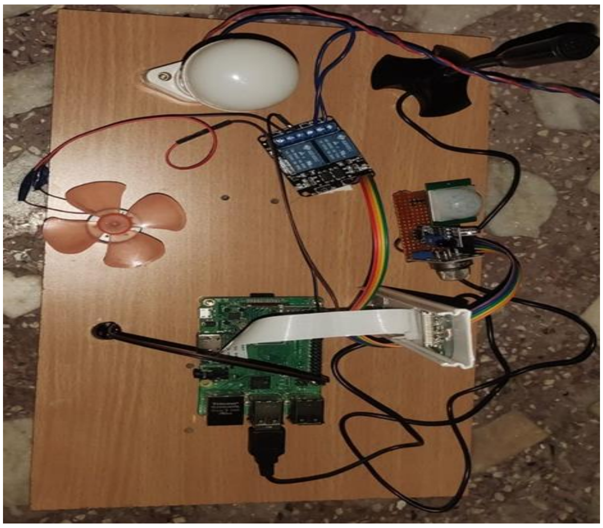


Fig : Hardware Setup

3 EXPERIMENTAL METHODOLOGY

This section explains the experimental setup, testing procedures, and validation process for the Gesture-Based & Voice (AI-Offline) Based Home Automation System. The methodology focuses on testing the functionality, accuracy, and performance of various modules within the system.

3.1 Hardware Testing

The hardware testing involves verifying the connectivity and functionality of all physical components used in the system. The process includes:

- Testing Raspberry Pi 3:

o Ensuring the GPIO pins are properly connected to the relay modules and sensors.

o Verifying power supply and functionality of the Raspberry Pi.

- Testing Camera Module (Gesture Recognition):

o Ensuring proper connection and functionality of the camera module.

o Testing image capture, preprocessing, and gesture detection using MediaPipe.

- Testing Microphone (Voice Recognition):

o Ensuring the microphone is correctly connected to the Raspberry Pi.

o Testing audio capture and voice command processing using Vosk AI.

- Testing Sensors (Motion, Gas, Fire/Smoke):

o Ensuring correct interfacing of sensors with the Raspberry Pi.

o Testing sensor sensitivity and response times under various environmental conditions.

- Testing Relay Modules:

o Verifying the switching mechanism to control devices (fan, light, etc.) based on input signals from the Raspberry Pi.

3.2 Software Testing

The software testing focuses on the gesture and voice recognition modules along with system integration. The process includes:

• Gesture Recognition Testing (MediaPipe):

o Testing various predefined gestures under different lighting conditions and backgrounds.

o Measuring accuracy and response time for detecting gestures.

o Comparing recognized gestures with predefined gestures for appropriate control signals.

• Voice Recognition Testing (Vosk AI):

o Testing the recognition of predefined commands (e.g., "Turn on the light", "Turn off the fan").

o Testing under various noise levels to assess performance in noisy environments.

o Measuring response time and accuracy of the voice recognition system.

3.3 Integration Testing

The integration testing ensures seamless communication between hardware and software components. This involves:

• Combining the gesture and voice recognition modules into a single Python-based application.

• Testing the simultaneous functioning of both modules without interference.

• Ensuring proper communication between the Raspberry Pi and relay modules for controlling appliances.

• Testing the Telegram Bot for sending alerts and receiving acknowledgments.

3.4 Performance Analysis

The performance of the system is analyzed based on the following metrics:

• Gesture Recognition Accuracy: Comparison of predicted gestures with predefined gestures.

• Voice Recognition Accuracy: Accuracy of detected commands under varying conditions.

• Response Time: Time taken from input detection to output execution (control signal generation).

• Reliability: Consistency of system performance over extended usage periods.

• Energy Efficiency: Effectiveness of the energy optimization mechanism based on motion detection.

4 RESULTS

The proposed Gesture-Based & Voice (AI-Offline) Based Home Automation System was tested extensively to evaluate its performance under various conditions. The results obtained are summarized as follows:

4.1 Gesture Recognition (MediaPipe)

• The gesture recognition system was tested with various predefined gestures such as palm open, fist, swipe left, and swipe right.

• The accuracy of gesture detection was approximately 90% under normal lighting conditions and 60% under low-light conditions.

• The average response time for gesture recognition was measured to be 1-2 seconds, which is sufficient for real-time control.

• We have used '1' as an input for switching on light and '3' as the input for switching on fan as well as '2' as an input for switching off light and '4' as an input for switching off fan.

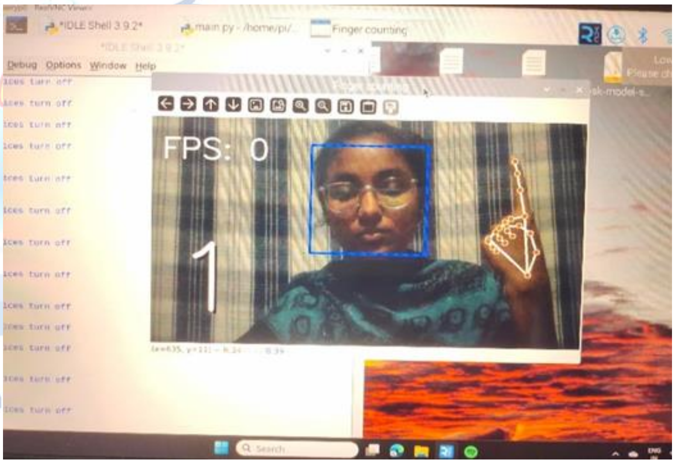


Fig: Gesture Input for switching on light

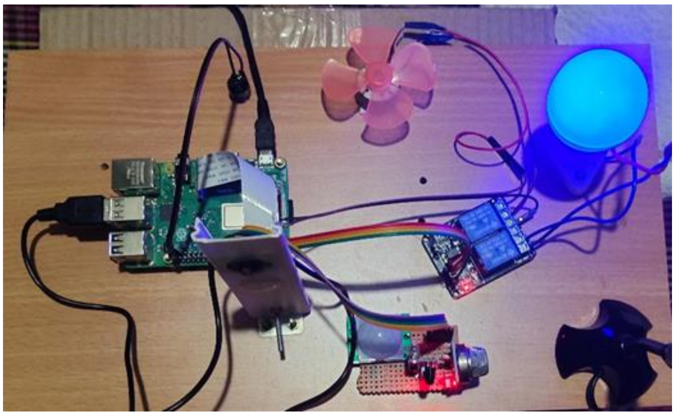


Fig: Output for the gesture

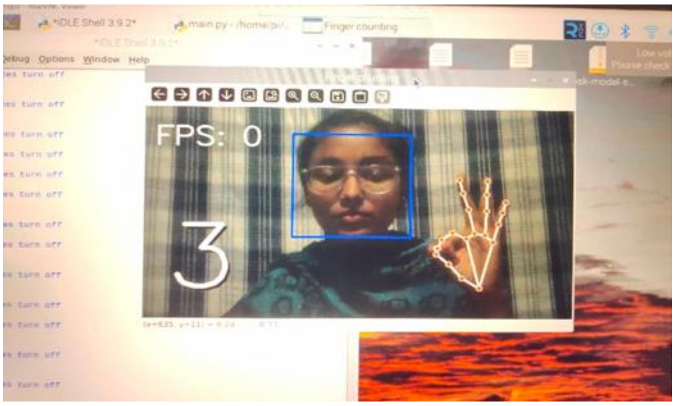


Fig: Gesture input for switching on fan

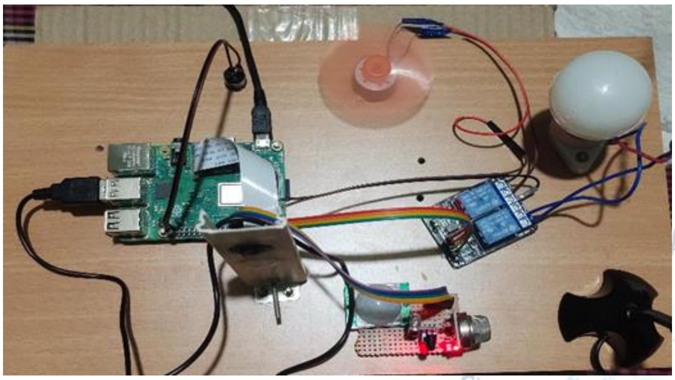


Fig : Output for the gesture

4.2 Voice Recognition (Vosk AI)

• The voice recognition system was tested with predefined commands: "hey", "hello ", to turn on and off the fan . "Light on", and "Light off" for light.

• The accuracy of command recognition was approximately 80% in quiet environments and 40% in noisy environments.

• The average response time for voice recognition was recorded as 2-3 seconds.

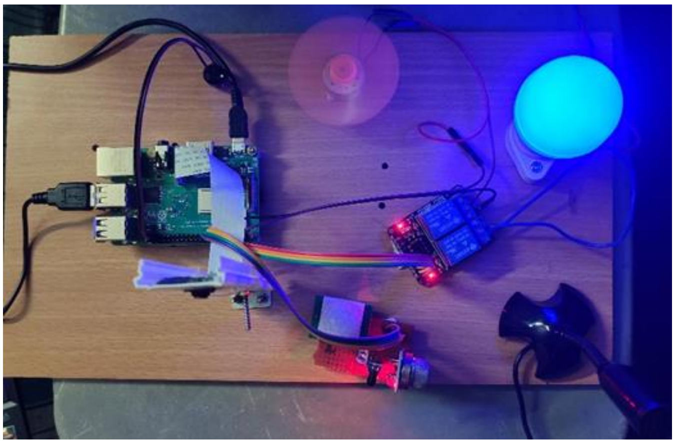


Fig: Output for the Voice commands

4.3 Sensor Performance

• The motion sensor effectively detected human presence within a range of 7-8 meters.

• The gas sensor successfully detected the presence of harmful gases and triggered appropriate alerts.

• The fire/smoke sensor provided accurate detection of smoke/fire with a response time of 3-4 seconds.

4.4 Energy Optimization

• The system successfully turned off appliances when no motion was detected for a predefined duration of 1 minute.

• Before implementing our system, appliances were often left running even when not in use, leading to excessive energy consumption. We calculated the total power usage by multiplying the wattage of each device by its daily operating hours.

4.5 System Reliability

• The overall system demonstrated stable and reliable performance over continuous operation for several hours.

• The integration of gesture and voice control modules was achieved without interference or delays.

5 CONCLUSIONS

The Gesture-Based & Voice (AI-Offline) Based Home Automation System was designed and implemented to provide an efficient, reliable, and user-friendly approach to controlling household appliances. The system integrates gesture recognition using MediaPipe, offline voice recognition using Vosk AI, and various sensors to enhance safety and energy efficiency. The Raspberry Pi serves as the central control unit, managing inputs from the camera module, microphone, and sensors to generate appropriate control signals for devices connected through relay modules. Additionally, a Telegram Bot was used for real-time alert notifications, ensuring effective monitoring and communication.

The results obtained from extensive testing of the system demonstrate the following:

• The gesture recognition module achieved a high accuracy rate under standard conditions and performed effectively in real-time.

• The offline voice recognition module successfully recognized predefined commands with reasonable accuracy, even in noisy environments.

• The sensors (motion, gas, and fire/smoke) accurately detected environmental conditions and triggered appropriate alerts via the Telegram Bot.

• The energy optimization mechanism effectively reduced unnecessary power consumption by turning off appliances when no motion was detected for a specific duration.

• The system provided a reliable and stable performance over extended periods of operation.

• The proposed system successfully integrates gesture recognition, offline voice control, and sensor-based monitoring to achieve an efficient and accessible home automation solution. The dual-mode control mechanism offers flexibility to users, particularly those with physical disabilities or limited internet access. The integration of sensors and a Telegram Bot ensures a high level of safety and real-time alert mechanisms. The energy optimization feature further enhances the system's overall efficiency.

• While the system achieved its primary objectives, future improvements could include expanding the number of gestures and voice commands, improving the robustness of the system under varying environmental conditions, and enhancing the hardware setup for better accuracy and response time.

6 FUTURE WORK

The proposed Gesture-Based & Voice (AI-Offline) Based Home Automation System provides a reliable and efficient method for controlling home appliances. However, there are several areas where improvements and enhancements can be made. Potential future work includes:

• Expanding the Gesture and Voice Command Library:

o Adding more predefined gestures and voice commands to provide greater control over various appliances.

o Implementing dynamic gesture recognition for user-defined gestures.

• Improving Gesture and Voice Recognition Accuracy:

o Enhancing the accuracy of gesture recognition under varying lighting conditions and complex backgrounds.

o Incorporating noise cancellation techniques to improve voice recognition in noisy environments.

• Integration with More Smart Devices:

o Expanding the system to control additional devices such as smart TVs, air conditioners, and security cameras.

o Supporting wireless communication protocols like Wi-Fi, Bluetooth, and Zigbee for broader device compatibility.

• Enhanced Security Features:

o Implementing user authentication mechanisms to prevent unauthorized access to the system.

o Adding secure communication channels for data transmission.

• Hardware Optimization:

o Upgrading to more powerful microcontrollers or processors to improve processing speed and efficiency.

o Utilizing low-power components to further optimize energy consumption.

• Mobile Application Development:

o Creating a dedicated mobile application for monitoring and controlling the system remotely.

o Allowing users to configure gestures and voice commands through an intuitive interface

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1]https://www.prnewswire.com/in/news-releases/factory-automation market-size-to-reach-usd-368-372-4-million-by-2025-valuates reports-898630302.html.

[2] https://www.fortunebusinessinsights.com/enquiry/request-sample pdf/home-automation-market-100074.

[3] Piyare, R., & Tazil, M. (2011). Bluetooth based home automation system using cell phone. 2011 IEEE 15th International Symposium on Consumer Electronics

[4] AlShu'eili, H., Gupta, G. S., & Mukhopadhyay, S. (2011). Voice recognition based wireless home automation system. 2011 4th International Conference on Mechatronics (ICOM)

[5] Arathi, P. N., Arthika, S., Ponmithra, S., Srinivasan, K., & Rukkumani, V. (2017). Gesture based home automation system. 2017 International Conference on Nextgen Electronic Technologies: Silicon to Software (ICNETS2).

[6]https://www.geeetech.com/wiki/index.php/Arduino_Voice_Recogni ti on_Module

[7]https://www.seeedstudio.com/blog/2019/11/21/nrf24l01-getting started-arduino-guide/

[8]https://lastminuteengineers.com/nrf24l01-arduino-wireless communication/