Int.J.HIT.TRANSC:ECCN.Vol.13: Special Issue NSEFCCIC (2026) Page 49-53 ISSN: 0973-6875

Available Online at www.hithaldia.in/locate/ECCN

Smart Home Automation and Security System Using Gesture Recognition, Voice Assistance, IoT, and Machine Learning

[1]Anik Kapat, [1]Aditya Kumar, [1]Honey Yadav, [1]Nanda Kishor Manna, [1]Muskan Sureka,

[2] Tilak Mukherjee

[1]UG student, Department of Electronics of Communication Engineering Engineering, Haldia Institute of Technology, Haldia, Purba Medinipur, West Bengal
[2]Department of Electronics of Communication Engineering Engineering, Haldia Instituteof Technology, Haldia, Purba Medinipur, West Bengal

ABSTRACT

In order to offer a touchless and intelligent home control solution, the suggested Smart Home Automation and Security System combines gesture recognition, voice assistance, IoT connectivity, and machine learning. With an ESP32-based architecture, the system allows users to operate appliances with voice requests and hand gestures. Motion detection and facial recognition work together to improve security by capturing evidences during illicit entry and generating alarms. Environmental sensors automate appliance functioning by continuously monitoring temperature and humidity. The system is a scalable and affordable option for contemporary smart homes since it offers real-time monitoring, control, and security upgrades via a centralized web and mobile dashboard.

KEYWORDS: Gesture Recognition; Voice Assistant; Internet of Things (IoT); Machine Learning; Home Automation; Smart Security

1. INTRODUCTION

Home management systems have advanced far beyond basic remote-controlled switching in the modern world due to the increased emphasis on accessibility, convenience, and residential safety. By physically pressing switches, twisting dials, or touching touch screens, the traditional approach of engaging with a home implicitly assumes that the user is present, able-bodied, and unoccupied. In a surprisingly large number of real-world scenarios, this assumption is incorrect. Elderly people might not have the mobility or dexterity to reliably operate ordinary controls. Touch-based interfaces may be completely inaccessible to people with physical disabilities. A system that responds to a gesture or a spoken word instead of requiring a free hand is beneficial even for a fully capable user who is cooking, grocery shopping, or caring for a child. The development of touch-free control mechanisms, particularly hand gesture recognition and voice command processing, which make the home environment more responsive to its occupants rather than requiring adaption from them, is motivated by these practical gaps. The possibilities for home automation have fundamentally changed due to the Internet of Things. IoT architectures make it possible to monitor and control a home from any location at any time using any networked device by connecting household equipment to a shared network and exposing them thru a common communication layer. A caregiver can validate that a medical gadget in a patient's room is functioning properly, and a user traveling overseas can check if a fan was left running. Once restricted to costly proprietary installations, this remote visibility and control capability is now possible with inexpensive microcontrollers and cloud messaging services, making it suitable for routine household deployment [1, 2].

The picture is further enhanced using machine learning. A machine-learning-powered system can identify patterns, such as a particular hand configuration, a spoken phrase in natural language, a familiar face, or an unusual temperature reading, and react to them meaningfully instead of requiring the user to interact with a fixed menu of discrete commands. As a result, the interface becomes more adaptable and flexible. It is possible to expand a gesture vocabulary without rebuilding the hardware. Without changing the underlying control logic, a voice model can be retrained to identify a new command [3]. Once restricted to costly proprietary installations, this remote visibility and control capability is now possible with inexpensive microcontrollers and cloud messaging services, making it suitable for routine household deployment.

It is impossible for a PIR sensor to determine whether someone crossing the threshold is a resident. This restriction results in frequent false alarms that are set off by pets, passing cars, or regular home movement. Over time, this erodes user confidence in the system and frequently causes alerts to be disregarded or disabled. This flaw is directly addressed by combining motion sensing and facial recognition: the system first determines that movement has taken place before determining if the responsible individual is identified. The system only escalates-disseminating an alert, setting off an alarm, and taking pictures as proof-when the face cannot be matched to an authorized occupant. This multilayered strategy maintains a dependable barrier against actual incursion while drastically reducing false positives [4, 5].

The capability set is completed with environmental awareness. Among the most direct factors influencing indoor comfort and, in some situations, are temperature and humidity. Unless the system can recognize these variables and react, a home that becomes uncomfortably hot in the afternoon or excessively humid during the monsoon months requires continuous human attention. The backend can automatically turn on a fan, air conditioner, or dehumidifier without the occupant's input by reading from an environmental sensor and comparing measurements against user-configured thresholds. This reduces discomfort and the mental strain of constantly monitoring the home environment [2].

All of these threads are combined into a single, cohesive platform by this initiative. Hands-free appliance control is made possible by voice interaction and gesture detection. Device communication and switching are managed by an IoT-capable microcontroller. Access control is enforced with few false alarms when face recognition and motion sensing are integrated. Automated climate reactions are powered by environmental sensors. Additionally, a single web and mobile interface connects everything, offering environmental readings, security alerts, and real-time status updates while maintaining the user's ability to manually override any automated decision at any moment. The system's components, operation, and experimental validation are all thoroughly explained in the sections that follow.

2. LITERATURE REVIEW

Recent advancements in the Internet of Things (IoT) have significantly transformed home automation systems by enabling remote monitoring and control of household appliances. Traditional IoT-based home automation systems primarily focus on smart phone-controlled switching and environmental monitoring. While these systems improve convenience, they often rely heavily on manual interaction and provide limited accessibility for elderly and physically challenged users [1, 6].

Gesture recognition technology has emerged as a promising alternative to conventional input methods. Computer vision and machine learning techniques allow users to control devices through natural hand movements without physical contact. However, most gesture-based systems focus solely on appliance control and do not integrate security mechanisms [3].

Voice-controlled home automation systems further enhance user convenience by enabling hands-free interaction through speech commands. Modern speech recognition technologies have achieved high accuracy, but voice-only systems may face challenges in noisy environments and often lack robust security features [4]. Similarly, many smart security systems employ motion sensors or surveillance cameras independently. PIR-based motion detection systems frequently generate false alarms because they cannot distinguish between authorized and unauthorized individuals. Face recognition technology addresses this limitation by providing identity verification capabilities [5].

The proposed system combines gesture recognition, voice assistance, IoT connectivity, environmental monitoring, motion detection, and face recognition into a unified platform. This integration enhances accessibility, security, and automation while reducing the limitations associated with standalone solutions [1-6].

3. SYSTEM ARCHITECTURE

3.1 Input Layer

The input layer is in charge of gathering environmental data and user commands. It has a camera for facial recognition and gesture recognition, a microphone for spoken command acquisition, a PIR sensor for motion detection, and a DHT11 sensor for temperature and humidity monitoring. Data is continuously collected by these devices and sent to the processing layer for analysis [2, 3, 5].

3.2 Processing and Communication Layer

Security validation, environmental analysis, speech processing, and gesture recognition are all carried out by the processing layer. Computer vision algorithms analyze camera frames to recognize allowed faces and predetermined hand motions. Spoken commands are translated into machine-readable instructions by a speech-to-text engine. All decision-making is coordinated by a centralized backend that uses MQTT-based communications to interact with the ESP32 microcontroller [1, 4].

3.3 Device Control Layer

The ESP32 microcontroller, relay modules, alarm unit, and linked appliances make up the device-control layer. The ESP32 reports the status of applications to the dashboard and activates or deactivates them in response to orders received from the backend. Additionally, this layer controls environmental automation and security alerts [6].

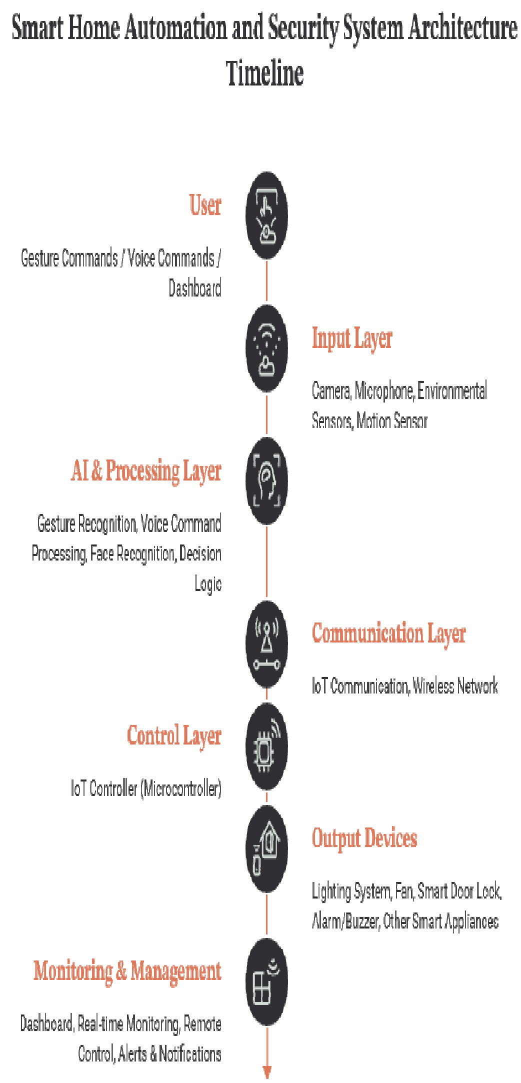

Figure 1: Smart home automation based architecture timeline

4. METHODOLOGY

4.1 Gesture and Voice-Based Appliance Control

The technology offers two ways to interact: voice instructions and hand gestures. A gesture-recognition module processes camera frames to classify predetermined hand configurations into commands for controlling appliances. Concurrently, a speech-recognition engine translates spoken commands into text and associates them with appropriate appliance operations. After being verified, both command streams are sent to the backend for processing [3, 4].

4.2 IoT Communication and Device Switching

The backend uses MQTT communication to send the instruction to the ESP32 microcontroller once a valid command has been created. After determining the proper relay channel, the ESP32 adjusts the appliance's state. To ensure real-time synchronization, updated device status is instantly sent back to the dashboard [1, 6].

4.3 Smart Security Monitoring

A PIR motion sensor and face-recognition module work together to provide security monitoring. Using the camera stream, motion detection initiates facial analysis. The event is recorded without taking any further action if the observed person fits the profile of an authorized user. The system sounds an alarm, takes pictures as proof, and notifies the user if the face cannot be identified or is missing [5].

4.4 Environmental Monitoring and Automation

The DHT11 sensor's temperature and humidity data are routinely compared to user-specified criteria. The backend automatically turns on appropriate appliances, like fans or air conditioners, when environmental conditions surpass predetermined thresholds. The mechanism returns the appliances to their original state once readings are within acceptable bounds [2].

5. PROPOSED IMPLEMENTATION OUTCOMES

It is anticipated that the suggested Smart Home Automation and It is anticipated that the suggested Smart Home Automation and Security System will offer a It is anticipated that the suggested Smart Home Automation and Security System will offer a smooth and user-friendly way to operate household equipment using voice and gesture commands. Reliable remote monitoring and device management via a centralized dashboard are anticipated to be made possible by the integration of IoT connection.

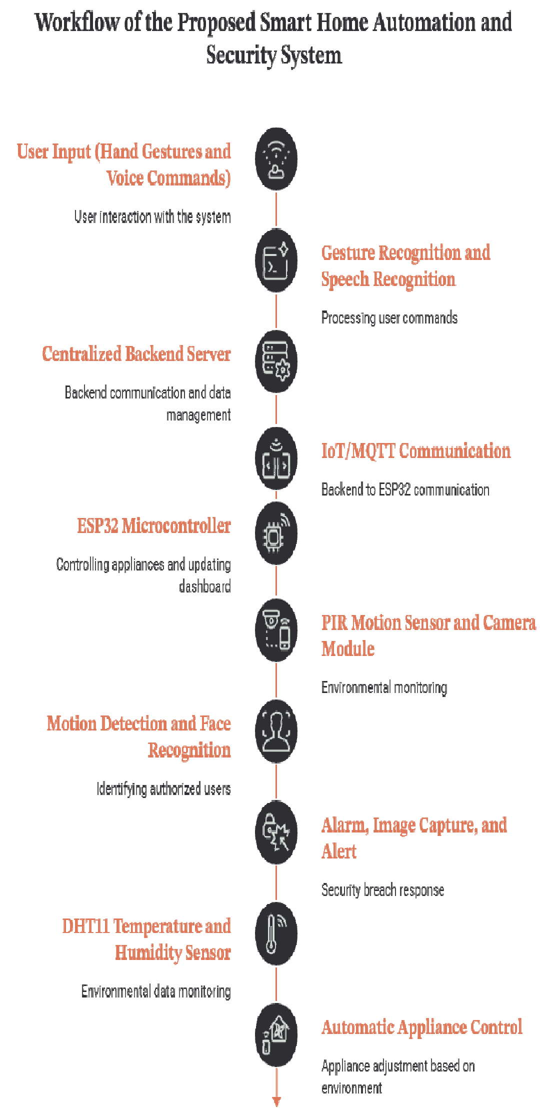

Figure 2: Workflow diagram of the proposed smart home secure automation system

By lowering the false alarms frequently connected to traditional motion-detection systems, the combined application of PIR-based motion sensing and facial recognition is anticipated to increase security. By automatically controlling linked appliances based on predetermined thresholds, environmental monitoring via temperature and humidity sensors is expected to improve user comfort.

Overall, it is anticipated that the suggested system will enhance domestic security, energy efficiency, accessibility, and convenience while preserving a modular architecture that can be expanded in subsequent deployments.

6. CONCLUSION

The proposed system demonstrates how voice interaction, gesture recognition, IoT connectivity, and machine learning can be combined into a unified, modular home-management platform that addresses security, accessibility, and convenience as a whole rather than as separate add-ons. The integrated use of motion sensing and face recognition for access control, combined with environment-aware automation and a centralized dashboard, shows the technique's potential for assistive technology, elder care, and general residential and small-commercial application. Future studies shall focus on expanding the platform to incorporate more sensors and appliances, improving identification accuracy in various illumination and noise settings, and expanding the language of gestures and voice instruction

REFERENCES

[1] Sivapriyan, R., Rao, K. M., & Harijyothi, M. (2020). Literature review of IoT based home automation system., Fourth International Conference on Inventive Systems and Control (ICISC), pp.101–105.

[2] Maswadi, K., Ghani, N. A., Hamid, S., & Rasheed, M. B. (2020). Systematic literature review of smart home monitoring technologies based on IoT for the elderly. IEEE Access, 8, 92244–92261. doi:10.1109/ACCESS.2020.2992727

[3] Sen, A., Mishra, T. K., & Dash, R. (2024). Novel human machine interface via robust hand gesture recognition system using channel pruned YOLOv5s model. arXiv preprint arXiv:2407.02585.

[4] Saini, S. K., Singh, A., & Singh, J. (2023). Review on voice controlled home automation system. International Journal of Engineering Research & Technology (IJERT), 12(5).

[5] Sruthy, S., Yamuna, S., & George, S. N. (2020). An IoT based active building surveillance system using Raspberry Pi and NodeMCU. arXiv preprint arXiv:2001.11340.

[6] Kabir, M. H., & Ahmed, M. T. (2020). IoT based low-cost smart home automation and security system using wireless technology. International Journal of Wireless and Microwave Technologies.