International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

GESTURE CONTROLLED SMART HOME USING ARDUINO AND DEEP LEARNING TECHNIQUES

S Satish Kumar[1*],S Sudheer Kumar [2], K Padma Sainath[3], K Swathi[4] , MD Shahid Rehan[5]

[1]Assistant Professor, [2]UG Student, [1,2]Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning)
[1,2] JB Institute of Engineering and Technology (UGC-Autonomous), Yenkapally, Moinabad, Hyderabad, 500075, Telangana.
*Corresponding author: S Sudheer Kumar (sudheernandini2004@gmail.com)

ABSTRACT

The rapid advancement of the Internet of Things (IoT) and Artificial Intelligence (AI) has significantly transformed traditional home environments into intelligent and automated systems. This paper presents the design and implementation of a Gesture Controlled Smart Home Automation System that integrates deep learning techniques with IoT-based device control to enhance user convenience and operational efficiency. The proposed system enables users to interact with household appliances through multiple control mechanisms, including hand gestures, voice commands, and a web-based interface. A vision-based gesture recognition module is developed using OpenCV for image acquisition and preprocessing, while a Convolutional Neural Network (CNN) model is utilized to accurately classify hand gestures in real time. The recognized commands are transmitted via wireless communication to an ESP32 microcontroller, which manages appliance operation through relay modules. Furthermore, a responsive web dashboard developed using modern web technologies provides remote monitoring and control capabilities. The system demonstrates reliable performance, improved accessibility, and reduced human effort, making it a practical and scalable solution for modern smart home environments. he integration of multiple control methods enhances system flexibility and usability while ensuring real-time response and reliable performance. The proposed system demonstrates improved accessibility, reduced manual effort, enhanced safety, and energy-efficient operation, making it a scalable and cost-effective solution for next-generation smart home environments and assistive technologies.

Key Words: Smart Home Automation, Internet of Things (IoT), Gesture Recognition, Deep Learning, Convolutional Neural Network (CNN), ESP32, Computer Vision, Web Dashboard.

1. INTRODUCTION

The rapid advancement of modern technologies such as the Internet of Things (IoT), Artificial Intelligence (AI), and wireless communication has significantly transformed traditional homes into intelligent and automated environments. Smart home automation systems have become increasingly popular due to their ability to enhance convenience, improve security, and optimize energy consumption. Conventional home control methods primarily depend on manual switches and direct physical interaction, which can limit flexibility, accessibility, and remote operation. As the demand for smart and efficient living environments continues to grow, the development of intelligent automation systems that support advanced and user-friendly interaction methods has become essential in modern society.

In recent years, gesture recognition and voice-based control technologies have gained significant attention due to their ability to provide contactless and intuitive control of household appliances. These technologies enable users to interact with home systems in a natural manner without the need for physical switches or remote controllers. Gesture recognition systems utilize computer vision techniques to identify human hand movements, while voice recognition systems allow users to issue commands through speech. Additionally, web-based control interfaces enable remote monitoring and management of home appliances through internet connectivity, allowing users to control devices from different locations. The integration of these technologies creates a flexible and efficient automation system capable of improving overall user experience and operational efficiency.

It presents the design and implementation of a Gesture Controlled Smart Home Automation System that integrates computer vision, deep learning, and IoT technologies to provide an intelligent and reliable home automation solution. The system uses a camera to capture hand gestures from the user, which are processed using image processing techniques implemented through OpenCV. A Convolutional Neural Network (CNN) model is employed to recognize predefined gestures accurately and convert them into corresponding control commands. These commands are transmitted wirelessly through a Wi-Fi network to an ESP32 microcontroller, which functions as the central control unit responsible for managing connected electrical appliances through relay modules.

In addition to gesture-based control, the system also supports voice command functionality that enables users to operate appliances using spoken instructions. Furthermore, a web-based dashboard is developed to provide users with the ability to monitor device status and control appliances remotely through a graphical user interface. The integration of gesture recognition, voice control, and web-based monitoring ensures improved system flexibility, enhanced usability, and efficient device management. The proposed system aims to deliver a cost-effective, scalable, and user-friendly smart home automation solution capable of meeting the evolving requirements of modern households while contributing to the advancement of intelligent home technologies. The implementation of wireless communication ensures reliable and real-time interaction between the user and home appliances, thereby improving system responsiveness and operational efficiency. By combining intelligent recognition techniques with IoT-based automation, the system contributes to the development of modern smart living environments.

2. LITERATURE SURVEY

Smart home automation has experienced significant growth in recent years due to the rapid advancement of the Internet of Things (IoT) and intelligent control technologies. Earlier home automation systems primarily relied on manual switches and basic remote-control mechanisms, which provided limited functionality and required direct physical interaction. These traditional approaches lacked flexibility, remote accessibility, and intelligent decision-making capabilities, making them less suitable for modern smart living environments. The introduction of IoT-based systems enabled devices to be connected through wireless networks, allowing users to monitor and control household appliances remotely using internet-enabled platforms.

One of the early developments in home automation involved the use of microcontroller-based systems that allowed users to control devices through mobile applications or SMS-based communication. These systems improved convenience and reduced manual effort; however, they often required dedicated hardware interfaces and lacked advanced user interaction features. With the advancement of wireless communication technologies such as Wi-Fi and Bluetooth, researchers developed more efficient systems capable of providing real-time control and monitoring of home appliances. Microcontrollers such as Arduino and ESP32 became widely used in smart home applications due to their low cost, energy efficiency, and built-in connectivity features.

To further improve user interaction, voice recognition technology was introduced into home automation systems. Voice-controlled systems allowed users to operate appliances using spoken commands, enabling hands-free control and improving accessibility for elderly and physically challenged individuals. Speech recognition algorithms were integrated with IoT devices to provide faster response times and enhanced usability. Although voice-based systems offered improved convenience, they sometimes faced challenges related to noise interference, language variations, and recognition accuracy in real-world environments.

Gesture recognition technology emerged as another important advancement in smart home automation, providing a natural and contactless method of controlling devices. Computer vision techniques were used to capture and analyze hand movements through cameras, enabling the system to interpret user gestures as control commands. The development of machine learning and deep learning algorithms, particularly Convolutional Neural Networks (CNN), significantly improved the accuracy and reliability of gesture recognition systems. These systems allowed users to control appliances without touching switches, making them suitable for environments where hygiene and safety are important considerations.

In addition to gesture and voice-based control methods, web-based monitoring and control systems have been widely implemented to provide remote accessibility and centralized device management. Web interfaces allow users to monitor the status of connected appliances and control them from any location using internet connectivity. The integration of IoT communication protocols ensures secure data transmission and efficient interaction between devices and users. However, many existing systems focus on a single control method, which may limit system flexibility and user convenience.

Recent research emphasizes the development of multi-modal smart home automation systems that combine multiple input methods within a single platform. By integrating gesture recognition, voice commands, and web-based control, these systems provide improved reliability, flexibility, and user experience. The proposed Gesture Controlled Smart Home Automation System builds upon these advancements by combining computer vision, deep learning, and IoT technologies to create an intelligent and efficient home automation solution capable of real-time device control and remote monitoring

3. PROPOSED SYSTEM

The proposed system presents an intelligent **Smart Home Automation System** designed to control household appliances using gesture recognition, voice commands, and a web-based interface through the integration of Internet of Things (IoT) technology and deep learning techniques. The system utilizes a camera to capture real-time hand gestures performed by the user, and computer vision techniques are applied using the OpenCV library to process and analyze the captured images. A Convolutional Neural Network (CNN) model is used to recognize predefined gestures and convert them into corresponding control commands for operating devices such as lights and fans. In addition to gesture recognition, the system incorporates a voice control mechanism that allows users to manage appliances through spoken instructions using speech recognition technology, enabling convenient hands-free interaction. Furthermore, the system includes a web-based interface that provides remote monitoring and control of connected devices through internet connectivity, allowing users to access and manage appliances from different locations. At the core of the system, an ESP32 microcontroller functions as the central control unit responsible for receiving commands from gesture, voice, and web modules and executing the required actions. The ESP32 communicates with electrical appliances through relay modules to safely regulate power supply and ensure reliable device operation. Wireless communication using Wi-Fi enables real-time

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

data transmission between system components, improving system responsiveness and efficiency. The proposed system is designed to be scalable, cost-effective, and user-friendly, supporting the integration of additional smart devices in the future while providing a reliable solution for modern smart home automation environments.

3.1 System Workflow

The workflow of the proposed Smart Home Automation System is designed to integrate three different input mechanisms—gesture recognition, voice commands, and a web-based interface—into a single unified control system that provides one consistent automation service. The system operation begins when the user provides an input through any one of the available control methods. In the gesture recognition module, the input is a hand gesture performed by the user in front of a camera. The camera captures the gesture image, which is then processed using image preprocessing techniques such as frame acquisition and feature extraction. The processed image is analyzed using a trained Convolutional Neural Network (CNN) model to identify the corresponding gesture command. Once the gesture is recognized, the system converts the detected gesture into a digital control signal.

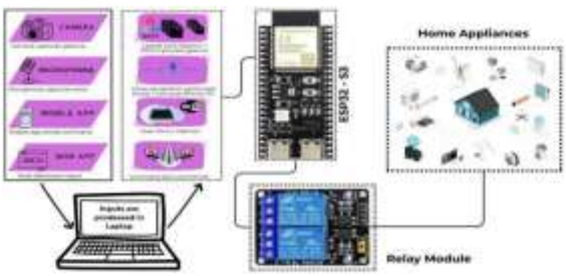

Figure 1. Proposed System Architecture of Gesture Control Smart Home

In the voice control module, the input is a spoken instruction provided by the user through a microphone. The microphone captures the voice signal, which is then processed using speech recognition algorithms to identify specific command words. The recognized voice command is converted into a digital instruction representing the required device operation. Similarly, in the web-based control module, the input is a user-selected command from a graphical interface accessed through a computer or mobile device. The command is transmitted over a wireless network or internet connection to the central system in real time. Each module independently generates a command signal based on the user input, ensuring flexibility in system interaction.

All command signals generated from the gesture, voice, and web modules are integrated and transmitted to the central controller, the ESP32 microcontroller, which acts as the core processing unit of the system. The ESP32 receives the command from any of the input sources and processes it to determine the appropriate control action. After validating the command, the microcontroller sends a control signal to the relay module, which functions as the output interface responsible for switching electrical appliances. The relay module controls the power supply to connected devices such as lights, fans, or other household appliances by turning them ON or OFF according to the received instruction. The final output of the system is the successful execution of the requested device operation. By integrating gesture recognition, voice commands, and web-based control into a single coordinated workflow, the system ensures reliable, flexible, and user-friendly smart home automation with multiple input options delivering one unified service.

analyzing the similarity between the text embedding and different regions of the image. Based on this analysis, the model predicts bounding boxes that indicate the location of detected objects. Each detection is associated with a confidence score representing the likelihood of a correct match. The inference process is efficient and supports real-time detection, and since the model is pre-trained, it can recognize a wide range of objects without additional training. Finally, the Output and Visualization Layer presents the results by overlaying bounding boxes on the original image and displaying confidence scores, making the output clear and easy to interpret for users.

Workflow ensures seamless integration of gesture recognition, voice commands, and web-based control into a single unified system for efficient device management. This integrated approach enables reliable communication between input modules and output devices, ensuring real-time response and consistent system performance.

4. RESULTS DESCRIPTION

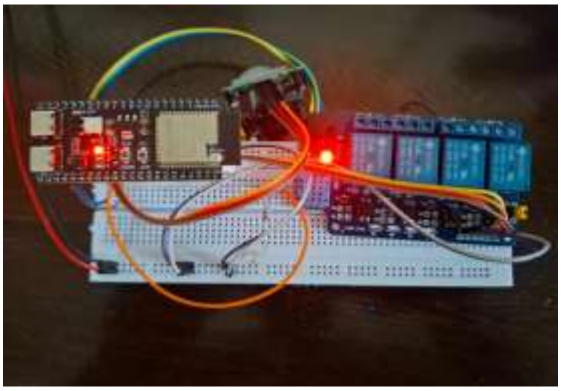

Figure 2. Hardware Module

The experimental hardware module integrates an ESP32-S3 microcontroller with a four-channel relay board (HW-316 driven) on a breadboard, enabling control of non-smart appliances via existing electrical wiring without requiring intrusive rewiring. Red LEDs on the ESP32 and relay indicate active power status, providing visual feedback on system operation. The relays act as switches for high-power loads (up to 10A 250VAC), allowing the ESP32 to toggle appliances based on voice, gesture, or web commands. This setup demonstrates a low-cost, retrofit solution for smart home automation, making legacy devices controllable through a unified interface. The configuration serves as a prototype foundation for sensor-integrated, AI-enabled control of domestic environments, with potential applications in energy monitoring, voice-controlled management, and IoT-based smart device prototyping. The modular design allows for easy expansion to control additional appliances or integrate with smart home ecosystems like Blynk or Home Assistant.

Figure 3. hardware code and compilation

The Arduino IDE screenshot shows a Blynk-enabled sketch for an ESP32-S3 Dev Module that manages four relays and an LED.The code uses Blynk.virtualWrite(vpin, state) to command virtual pins (V_RELAY1–V_RELAY4) linked to physical relays. A setRelay function toggles each relay based on Blynk parameters, storing states in relay1State…relay4State. It also controls an LED with digitalWrite(LED_PIN, state) and implements a system-lock mode via V_MODE.Compilation is ongoing on COM5, with progress reports at addresses like 0x0000… and memory usage shown as percentages. The firmware size is 58,7140 bytes, and the upload reaches 100 % completion.

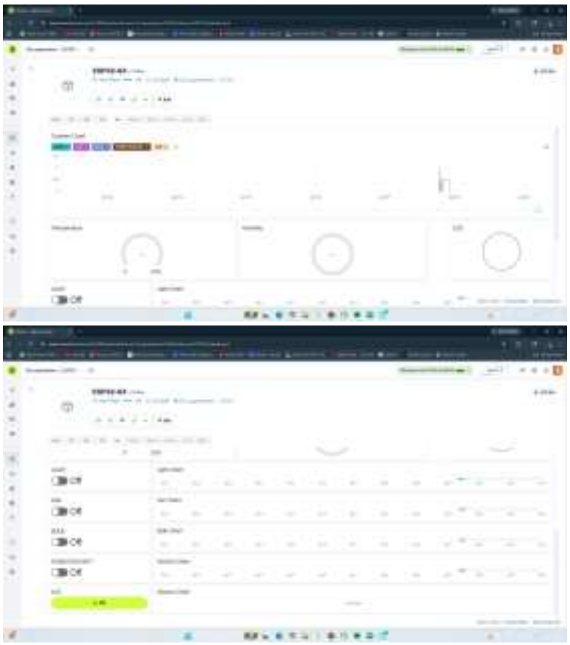

Figure 4,5.World wide web control Dash Board

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

The Blynk Console dashboard showcases a web-based control interface for an ESP32-S3 device, facilitating smart home automation. The device is online, associated with "My organization - 1113SY", and features toggle switches for controlling LIGHT, FAN, BULB, POWER SOCKET, and LED, all currently off. The interface includes data visualization tools like a Custom Chart and gauges for Temperature, Humidity, and LED status, with historical data spanning Apr 02 to Apr 09. Message usage is tracked, indicating 52.3k/100k and 52.4k/100k messages used across the screenshots. This setup enables worldwide remote control of appliances and monitoring of environmental parameters, suggesting integration with IoT platforms for Smart home automation.

https://drive.google.com/file/d/1dWri9OiFmku1Qt76iWHWuEqTpQrGu770/view?usp=drive_link

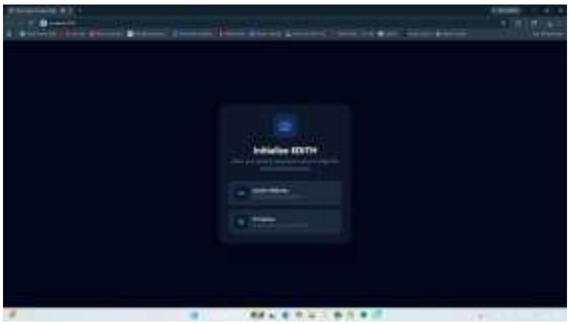

figure 6. initialization screen for EDITH

The image shows an initialization screen for *EDITH* (likely an AI or synchronization tool), prompting the user to select a primary visual input source to begin the synchronization process. The options are:

1. *System Webcam* – direct hardware integration.

2. *IP Camera* – network stream via RTSP/HTTP.

The interface is displayed in a browser window running on `localhost:3000`, titled "My Google AI Studio App".

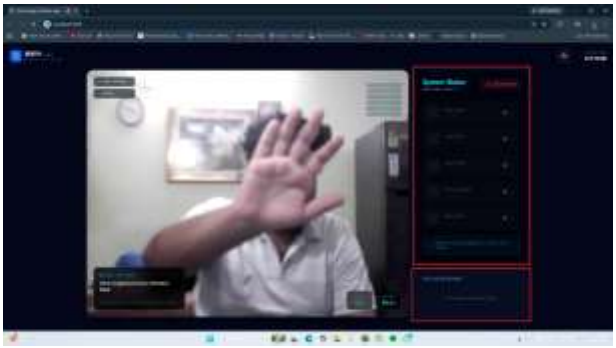

figure 7. Face not Recognized so Control Dashboard, Voice & Gesture Commands are Disabled. And we can see no active persons in Active Sessions.

The image shows an authentication error in an AI studio app where:

- Face verification failed – the system recognized the voice but could not verify the face, so the control dashboard, voice, and gesture commands are disabled.
- System Status is RESTRICTED because biometric verification (saying "EDITH" or using gestures) is required to unlock the controls.
- Active Sessions report "No subjects detected in frame", meaning the camera isn't detecting any person in the scene.

To resolve this, you need to perform the required biometric verification (say the activation phrase "EDITH" or use a recognized gesture) so the system can enable the controls and detect the user in the frame.

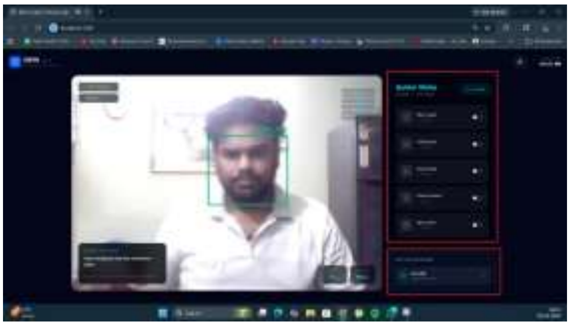

figure 8. Face and Voice authention complete. So all Kinds of Operations (Gesture, Web, & Voice) are active and can see person in Active Session.

The image shows a successful face and voice authentication process, with the system indicating that all operations (Gesture, Web, & Voice) are now active and the person is listed in Active Sessions with a verified identity (sai_nath).The green box highlights facial recognition with landmark dots, and the red boxes show the System Status (SECURE) and Active Sessions panels confirming the user is authenticated.The message at the bottom left reads “Voice recognized, but face verification failed,” which seems contradictory to the overall status shown (authentication complete).

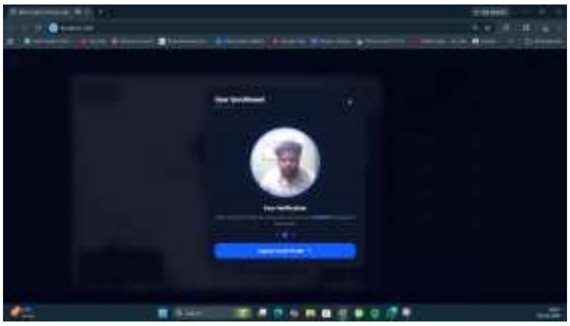

figure 9. Can add or remove Users for Control

Think of it like managing a crew in a game – you gotta go to the “admin” or “settings” panel where the boss (the system) lets you tweak who’s in the team.

1. Open the admin dashboard of your app (usually a “Users” or “Access Control” tab). There you’ll see buttons like “Add User” or “Delete User” – just click ’em and fill in the details.
2. If you’re coding it, you’ll use API calls like POST /users to add a new dude or DELETE /users/{id} to kick someone out.
3. Or straight in the database (if you have access), you’d run SQL queries to insert or delete user records – but be careful, you don’t wanna mess up the data.

4.1 Summary of Results :

- Admin UI – Go to the app’s admin/dashboard → “Users” section. Use the “Add User” button to create a new user or the “Delete” option to remove one.
- API method – Use HTTP requests: POST /api/users to add a user & DELETE /api/users/{id} to remove a user.
- Database – Run SQL queries: INSERT INTO users ... to add, or DELETE FROM users WHERE id = ... to remove

5.CONCLUSION:

This paper presents a comprehensive smart home automation system leveraging the ESP32-S3 microcontroller, Blynk IoT platform, and AI-powered voice/gesture control mechanisms. The proposed system seamlessly integrates legacy appliances into a unified ecosystem, enabling users to remotely monitor and control devices through a user-friendly Blynk dashboard. By incorporating face and voice authentication, the system ensures secure access, allowing only authorized users to manage appliances via voice commands, web interface, or gesture-based inputs. The modular design of the prototype facilitates easy expansion, making it an attractive solution for retrofitting existing homes with smart automation capabilities. The successful implementation of this system demonstrates a cost-effective approach to transforming conventional homes into intelligent, interconnected environments. Furthermore, the research lays the groundwork for future enhancements, including AI-driven predictive automation, energy optimization, and integration with broader smart home ecosystems, positioning this work at the forefront of smart home innovation and sustainable living solutions.

6.REFERNCES :

[1] Luigi Atzori, Antonio Iera, and Giacomo Morabito, “Internet of Things: A Vision, Architectural Elements, and Future Directions,” *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

[2] Rajkumar Buyya, Amir Vahid Dastjerdi, *Internet of Things: Principles and Paradigms*, Morgan Kaufmann, 2016.

[3] International Telecommunication Union, "Overview of the Internet of Things," ITU-T Recommendation Y.2060, 2012.

[4] Espressif Systems, *ESP32-S3 Technical Reference Manual*, 2023.

[5] Neil Kolban, *Kolban's Book on ESP32*, 2017.

[6] Daniel Jurafsky and James H. Martin, *Speech and Language Processing*, 3rd ed., Pearson, 2023.

[7] Google, "Speech Recognition with Deep Neural Networks," Google AI Research, 2022.

[8] OpenAI, "Advances in Natural Language Processing and Conversational AI," 2023.

[9] Gary Bradski and Adrian Kaehler, *Learning OpenCV: Computer Vision with OpenCV Library*, O'Reilly Media, 2008.

[10] OpenCV, "Open Source Computer Vision Library Documentation," 2024.

[11] IEEE, "Hand Gesture Recognition Using Computer Vision Techniques," IEEE Xplore, 2021.

[12] DHT11 Sensor, Datasheet, Aosong Electronics, 2020.

[13] PIR Motion Sensor, Technical Datasheet, 2021.

[14] Relay Module, Technical Specifications, 2022.

[15] Google Firebase, "Firebase Documentation," 2024.

[16] Amazon Web Services, "AWS IoT Core Developer Guide," 2023.

[17] Samsung, "SmartThings API Documentation," 2024.