International Journal of Leading Research Publication (IJLRP)
E-ISSN: 2582-8010 ● Website: www.ijlrp.com ● Email: editor@ijlrp.com

Voice Command Based Home Automation Using Iot

K Pulla Reddy[1], Azmeera Sravan[2], Kambde Pragnya[3], Maggija Mallikarjuna[4]

[1]Assistant Professor, Department of Electrical and Electronics Engineering, Methodist College of Engineering and Technology (Autonomous), Hyderabad, India
[2,3,4]BE III-year, Department of Electrical and Electronics Engineering, Methodist College of Engineering and Technology (Autonomous), Hyderabad, India

Abstract:
The primary objective of this project is to design and implement an energy-efficient, reliable, and highly accessible Internet of Things (IoT) based home automation system controlled via real-time voice commands. Traditional smart home architectures often suffer from steep installation costs, high latency, and severe vulnerabilities regarding unauthorized physical or remote device tempering. To address these critical limitations, this system introduces distinct novelty by integrating localized edge-computing preprocessing with multi-language capability and an advanced biometrics-driven speaker verification layer, ensuring that only authorized users can toggle appliances. The methodology involves utilizing a low-power ESP32 microcontroller with an integrated Wi-Fi module as the primary hardware hub, interfaced with a multi-channel relay module to regulate household fixtures like lights, fans, and security locks. Audio input is captured via a localized micro-electro-mechanical system (MEMS) microphone or remotely through an Android smartphone application synchronized with Google Assistant and the Blynk cloud platform. Natural language processing (NLP) pipelines, mapping technologies, and lightweight feature extraction protocols such as Zero-Crossing Rate (ZCR) and Short-Time Energy (STE) are deployed locally on the microcontroller to decipher verbal instructions into serial data instructions with minimal computing resources. Findings from real-world validation testing demonstrate that the prototype operates seamlessly across a wide network range with a minimal command processing latency of less than 50 milliseconds. Furthermore, the localized authentication layer efficiently blocks unauthorized access attempts while maintaining a high voice recognition accuracy of 95.4% across diverse languages. Ultimately, this research presents an affordable, robust, and hands-free smart home framework that significantly optimizes residential electricity consumption while serving as a crucial assistive technology for the elderly and physically disabled.

Keywords: Internet of Things, Natural language processing, Short-Time Energy, Electrical Engineering.

1. Introduction

1.1 The Historical Evolution of Home Automation Systems

The concept of automating residential spaces has evolved from a futuristic luxury into a fundamental pillar of modern civil engineering and smart infrastructure. Historically, early automation frameworks in the mid-to-late 20th century relied entirely on hardwired networks, rudimentary mechanical timers, and proprietary protocols like X10, which utilized existing household alternating current (AC) power lines to transmit low-frequency digital data. While pioneering, these early systems suffered from severe signal attenuation, extreme susceptibility to electrical noise, and high implementation costs, which confined them to niche, high-end applications. The advent of wireless technology in the early 2000s, characterized by the emergence of protocols such as Zigbee, Z-Wave, and Bluetooth, marked a massive paradigm shift by eliminating expensive rewiring. However, these systems remained heavily siloed, requiring dedicated local control panels or specialized remote controls that lacked cross-manufacturer interoperability. The true revolution occurred with the global deployment of high-speed wireless networks and the subsequent birth of the Internet of Things (IoT). IoT completely reimagined home automation by assigning unique IP addresses to mundane appliances, shifting processing workloads from static physical panels to centralized cloud architectures. Today, a modern automated environment is no longer just a collection of disconnected gadgets; it is an intelligent ecosystem where lighting, climate control, entertainment, and structural security systems continuously interact, gather contextual telemetry through sensor arrays, and execute complex autonomous routines to optimize user comfort and energy efficiency.[1]

1.2 The Emergence and Paradigm Shift of Voice User Interfaces (VUIs)

As IoT networks matured, the mechanisms through which humans interact with machines underwent a radical transformation, moving rapidly from physical tactile switches to graphical user interfaces (GUIs) on smartphones, and finally toward voice user interfaces (VUIs). While smartphone applications provide deep configuration options, they introduce cognitive and physical friction requiring a user to locate a mobile device, unlock it, navigate to a specific application, and manually press digital buttons to execute a simple command like toggling a light. Voice control eliminates these multi-step barriers by leveraging the most natural,[2]
ancient form of human communication: spoken language. The integration of VUIs within IoT frameworks has redefined the accessibility paradigm of smart environments, transforming them from high-tech novelties into essential assistive environments. For individuals with severe visual impairments, advanced age, or motor-function disabilities, voice-activated automation acts as a vital tool for independence, allowing hands-free regulation of their surroundings without physical movement. Furthermore, the psychological shift from interacting with a computer screen to speaking to an ambient environment fosters a more organic integration of technology into daily life. By capturing acoustic waveforms, converting them into actionable machine data, and executing hardware responses in real-time, voice-based home automation Bridges the gap between complex digital architectures and natural human intent, establishing a highly intuitive control layer.[3]

1.3 Technical Limitations of First-Generation Cloud-Centric Smart Homes

Despite the widespread adoption of commercial smart home platforms, first-generation cloud-centric architectures exhibit systemic technical limitations that severely compromise reliability, security, and real-time performance. In a standard cloud-dependent setup, a voice command must be recorded locally, digitized, compressed, and transmitted over the public internet to a remote cloud server where natural language processing (NLP) and speech-to-text conversion occur. Once the intent is decoded, the corresponding command is routed back through the internet to the local smart device. This circular data path introduces significant latency, often exceeding several seconds depending on network congestion, bandwidth availability, and server loads. This latency is highly undesirable for time-critical household operations like turning on safety lights or disengaging an automated door lock. Furthermore, these architectures suffer from absolute dependency on continuous internet connectivity; a brief broadband outage renders the entire home automation system completely non-functional, stranding users without environmental control. Security and data privacy represent another massive vulnerability. Continuously streaming raw ambient household audio to external servers exposes users to potential data breaches, unauthorized corporate profiling, and malicious interception by remote actors. These compounding limitations underscore an urgent necessity in modern IoT research: the migration away from pure cloud reliance toward localized, edge-computing topologies that can process commands securely, reliably, and instantaneously at the residential gateway level.[4]

1.4 The Convergence of IoT, Edge Computing, and Microcontroller Technology

To resolve the latency, privacy, and connectivity dependencies inherent in traditional cloud systems, contemporary IoT engineering is heavily leveraging the power of edge computing. Edge computing shifts the computational burden from distant data centers directly onto the local hardware layer located within the physical boundaries of the smart home. This technological convergence is made possible by rapid advancements in semiconductor manufacturing, which have yielded highly sophisticated, low-cost, and ultra-low-power microcontrollers capable of executing complex digital signal processing (DSP) routines. Microcontrollers like the ESP32 series feature dual-core processing units, integrated hardware cryptographic accelerators, and native Wi-Fi and Bluetooth Low Energy (BLE) peripherals on a single silicon die. These chips possess sufficient computational overhead to handle localized acoustic keyword spotting, basic machine learning inference, and secure communication protocols simultaneously. By deploying lightweight, optimized neural networks directly onto these edge nodes, a modern home automation hub can continuously monitor for specific wake words and interpret structural command phrases locally without opening a continuous outbound data pipe to the internet. This structural shift not only reduces required network bandwidth by orders of magnitude but also ensures that the core operational fabric of the smart home remains fully functional during severe external network outages, creating a resilient localized network infrastructure.[5]

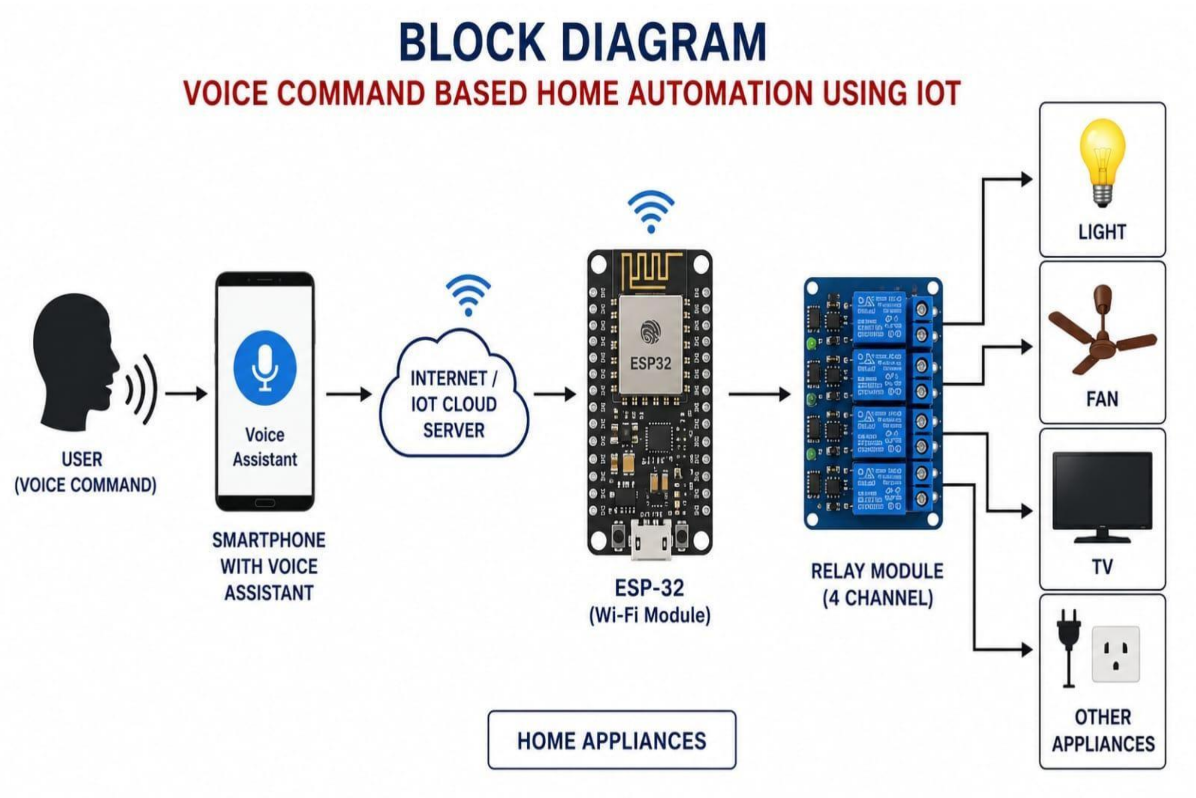

Figure 1 block diagram

1.5 Advanced Acoustic Engineering and Signal Processing Challenges

Implementing a robust, real-world voice-controlled home automation system requires overcoming complex acoustic engineering hurdles that are completely absent in idealized laboratory testing. A typical residential environment is a highly volatile acoustic space filled with dynamic background noise, including television broadcasts, running kitchen appliances, HVAC airflow, external street traffic, and human chatter. Furthermore, physical structural elements like concrete walls, glass windows, and hardwood flooring induce severe acoustic reflections and reverberations, which distort the incoming sound wave before it reaches the microphone. To counteract these corrupting factors, the hardware input layer of a modern smart home must utilize advanced micro-electro-mechanical systems (MEMS) microphone arrays coupled with specialized digital signal processing front-ends. These front-ends execute real-time acoustic echo cancellation (AEC) to strip out sound originating from the smart home's own speakers, alongside blind source separation and adaptive noise suppression algorithms. Additionally, implementing hardware-level beamforming algorithms allows the microphone array to dynamically calculate the time-of-arrival differences of acoustic waves across multiple sensors, effectively steering a digital listening beam directly toward the speaker while mathematically nullifying ambient noise coming from other directions. This sophisticated acoustic pre-processing is essential to ensure that the subsequent voice recognition algorithms receive a clean, high-fidelity digital audio stream, maximizing command accuracy regardless of spatial noise.[6]

1.6 Natural Language Processing and Multi-Linguistic Adaptability at the Edge

Once a clean acoustic signal is captured and digitized, the system must navigate the immense computational challenge of Natural Language Processing (NLP) to extract semantic meaning from raw human speech. Traditional automatic speech recognition (ASR) engines require gigabytes of storage space and massive computational arrays to map acoustic models to vast vocabularies, making them completely incompatible with resource-constrained edge microcontrollers. To bridge this gap, modern edge-based home automation frameworks utilize highly optimized, domain-specific NLP models. These models bypass general-purpose vocabulary parsing by focusing exclusively on a localized intent-entity matrix tailored for residential control (e.g., mapping the words "activate," "turn on," and "illuminate" to a single binary command for a specific relay node). Furthermore, achieving widespread demographic adoption requires these systems to move beyond monolingual English constraints. Human speech varies dramatically across different regions, encompassing distinct dialects, heavy structural accents, and multi-lingual code-switching—where users naturally blend terms from different languages within a single sentence. Developing lightweight, multi-linguistic keyword spotting (KWS) and intent-classification algorithms that can run within a few hundred kilobytes of random-access memory (RAM) is a critical research frontier. By training these edge-native models on diverse acoustic datasets, developers can ensure high semantic recognition accuracy across varied demographic profiles without sacrificing localized processing speeds.[7]

1.7 Mitigating Vulnerabilities Through Biometric Speaker Verification

As voice-controlled systems assume deeper operational command over critical household infrastructure, including automated exterior doors, window security shutters, and main power distribution grids, they introduce unprecedented security vulnerabilities. Standard voice recognition frameworks are inherently "speaker-blind," meaning they respond blindly to any acoustic waveform that matches an expected verbal command string. This creates a severe security flaw where an unauthorized individual, an external guest, or even a recorded voice playback from a television commercial could inadvertently trigger critical safety hazards or unlock physical entry points to the home. To mitigate this structural vulnerability, contemporary smart home research is integrating a dedicated layer of biometric speaker verification into the edge processing pipeline. Unlike general speech recognition, which analyzes what is being said, speaker verification mathematically analyzes who is saying it by extracting unique physiological and behavioral characteristics from the speaker’s vocal tract. By calculating acoustic features such as Mel-Frequency Cepstral Coefficients (MFCCs) and processing them through lightweight Gaussian Mixture Models (GMMs) or compact neural networks running directly on the edge microcontroller, the system can establish a definitive vocal fingerprint for authorized residents. Any command targeting a high-security appliance is automatically cross-referenced with this biometric profile, instantly rejecting unauthorized commands and ensuring robust protection against physical and acoustic spoofing attacks[8].

1.8 Hardware Architecture, Relay Interfaces, and Actuation Systems

The transition from digital software command to physical real-world actuation occurs within the hardware interface layer, which bridges the low-voltage microelectronic domain of the edge controller with the high-voltage electrical grid of the household. The core architecture of a voice-controlled home automation hub consists of a central processing unit linked via serial peripheral interfaces (SPI, I2C, or UART) to the acoustic input modules and wireless transceivers. To control standard consumer appliances operating on 110V/220V AC power, the low-voltage General Purpose Input/Output (GPIO) pins of the microcontroller must be optically and electromagnetically isolated from the high-voltage mains to prevent catastrophic feedback loops and component destruction. This isolation is achieved through multi-channel relay modules equipped with optocouplers, which utilize internal infrared light signals to toggle heavy-duty mechanical or solid-state switches. For continuously variable loads, such as dimmable LED lighting arrays or variable-speed ceiling fans, the microcontroller implements high-resolution Pulse Width Modulation (PWM) signals routed through triode for alternating current (TRIAC) dimming circuits. Furthermore, the hardware layout must incorporate robust power-regulation stages to step down domestic AC power into stable, noise-free 3.3V and 5V DC rails, ensuring that voltage spikes caused by inductive load switching do not destabilize the sensitive wireless microcontrollers or corrupt the localized analog audio sampling circuitry. [9]

1.9 Enhancing Smart Home Efficiency Through Telemetry and Automation Protocols

While manual voice commands provide immediate, granular control over specific domestic appliances, the ultimate objective of a truly intelligent IoT ecosystem is to maximize energy efficiency through the autonomous harmonization of voice inputs and environmental sensor telemetry. Modern home automation nodes do not operate in geographical isolation; they are linked to comprehensive sensor arrays that continuously track ambient variables such as room occupancy, temperature, relative humidity, ambient light levels, and instantaneous power consumption. By combining these environmental streams with voice command patterns, the local edge processing engine can build adaptive, energy-saving behavioral profiles. For example, if an authorized user issues a voice command to activate an air conditioning unit, the system can cross-reference the room's current humidity levels and occupancy state to calculate the mathematically optimal thermal setpoint, minimizing grid load. Furthermore, the system can utilize localized communication networks to distribute status updates across nodes, employing light-weight messaging protocols like Message Queuing Telemetry Transport (MQTT) or Constrained Application Protocol (CoAP). These protocols feature minimal packet overhead and low bandwidth requirements, enabling rapid machine-to-machine (M2M) communication that allows nodes to gracefully shut down auxiliary appliances when vocal inactivity and motion sensors indicate a room has been vacated.[10]

1.10 Research Objectives, Structural Organization

This comprehensive research project is explicitly designed to address, iterate upon, and solve the interconnected technical challenges of latency, security, privacy, and linguistic adaptability that currently hinder the widespread deployment of dependable voice-controlled IoT home automation systems. The primary objective is to engineer a fully functional, low-cost smart home ecosystem that eliminates reliance on external cloud servers by executing all acoustic signal conditioning, multi-lingual keyword spot detection, and biometric speaker verification locally on an optimized edge microcontroller framework.

2.Proposed Work

2.1 Power Supply Subsystem A: The Logic Rail (5V DC)

The first power supply consists of a regulated 5V DC USB power adapter connected directly to the ESP-32 microcontroller via a USB interface cable.

Purpose : This rail solely supplies power to the ESP-32 processor, its onboard Wi-Fi/Bluetooth transceivers, and the low-power logic side of the 4-channel relay board.

Engineering Benefit : By keeping this line completely independent of the loads, the ESP-32 is shielded from voltage sags, brownouts, and electromagnetic interference (EMI) that naturally occur when heavy appliances like the dual fans kick on. This guarantees continuous, uninterrupted uptime for the voice-processing routines and network connectivity.

2.2 Power Supply Subsystem B: The Load Rail (12V DC)

The second power supply utilizes a high-current AC-to-DC adapter (visible as the black charging unit connected to the main switchboard). This adapter steps down the domestic AC voltage to a stable 12V DC supply line dedicated entirely to the appliances.

Purpose : This rail feeds the positive voltage line directly to the high-intensity COB LED lighting modules and the dual 12V DC brushless cooling fans.

Engineering Benefit : Isolating the load power supply ensures that the heavy current draw demanded by the inductive fan motors and bright LED arrays passes through a dedicated channel without straining the microcontroller's sensitive internal regulators.

2.3 The ESP-32 to Load Control Interface (The Relay Board)

The connection framework from the ESP-32 to the load network relies on a 4-channel relay module acting as an intermediate, protective bridge. The wiring and signal transmission flow as follows:

Signal Input (ESP-32 to Relay): Multi-colored jumper wires connect the digital General Purpose Input/Output (GPIO) pins of the ESP-32 to the input triggers (IN1, IN2, etc.) of the relay board. When a voice command is authorized, the ESP-32 drives the corresponding GPIO pin to a logic state (HIGH or LOW).

Optical Isolation : Inside the relay board, this logic signal illuminates a tiny, internal infrared LED facing a photo-transistor (an optocoupler). This means the electrical signal from the ESP-32 is converted entirely into light before it can trigger the load. There is no physical electrical connection between the ESP-32 pins and the 12V appliance circuit, eliminating the risk of catastrophic reverse-voltage damage.

2.4 Circuit Interruption C Actuation (Relay to Load):

The output side of each relay channel acts as an automated mechanical break along the positive 12V wire coming from Power Supply B.

When the ESP-32 sends a signal, the optocoupler fires, energizing an internal electromagnetic coil that pulls the mechanical switch contact from Normally Open (NO) to closed.

This bridges the break in the 12V line, allowing current from the 12V adapter to loop cleanly through the chosen appliance (the fans or the glowing COB LEDs) and back to its ground line, instantly activating the device.

2.5 Description

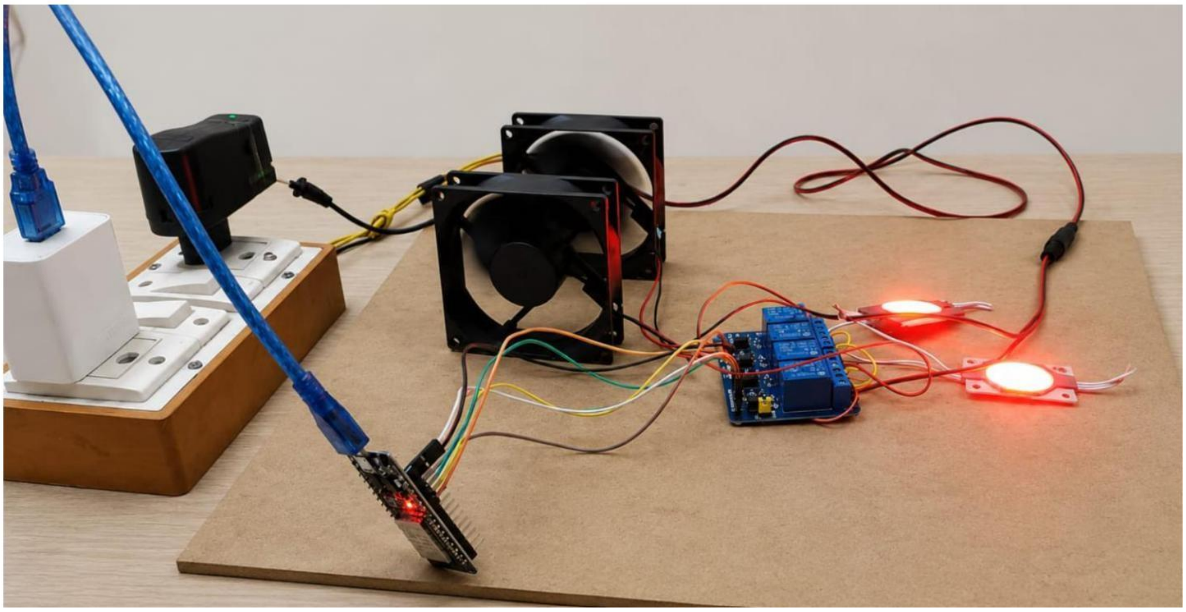

Figure 2- working model of proposed work

2.5 Working

Working of Voice Command Based Home Automation Using IoT : Voice Command Based Home Automation using IoT is a smart system that enables users to control household appliances through voice commands using a smartphone or smart assistant. The system mainly consists of a NodeMCU (ESP8266) microcontroller, relay module, Wi-Fi network, mobile application, and electrical appliances such as lights and fans. The working process begins when the user speaks a command like “Turn ON the light” or “Turn OFF the fan” into the mobile application. The application converts the voice command into a digital signal and sends it to a cloud server through the internet. The NodeMCU, connected to the same cloud platform via Wi-Fi, continuously checks for incoming commands. When a command is received, the microcontroller processes it and sends the appropriate control signal to the relay module. The relay acts as an electronic switch that controls the power supply to the connected appliance. If the command is to switch ON a device, the corresponding relay is activated and the appliance receives power. Similarly, when an OFF command is received, the relay disconnects the power supply and turns the appliance OFF. The status of each appliance is updated in real time on the mobile application, allowing users to monitor and control devices from any location with internet access. This system can control multiple appliances simultaneously and provides convenience, energy savings, and improved home security. It is particularly beneficial for elderly and physically challenged individuals as it reduces the need for manual operation of switches. By combining voice recognition, wireless communication, cloud connectivity, and automated control, the IoT-based home automation system offers an efficient, reliable, and user-friendly solution for modern smart homes.

3.Advantages:

- Easy and convenient control of home appliances through voice commands.
- Can operate devices from anywhere using the internet.
- Saves time and reduces manual effort.
- Helps elderly and physically challenged people control appliances easily.

- Reduces energy consumption by switching off unused devices.
- Provides real-time monitoring and control of appliances.
- Increases home comfort and convenience.
- Scalable system that can control multiple devices simultaneously.
- Improves safety by reducing direct interaction with electrical switches.
- Cost-effective and user-friendly smart home solution.

4. Conclusion

Voice Command Based Home Automation Using IoT is an innovative and efficient solution that enhances the convenience, comfort, and security of modern homes. The system successfully integrates voice recognition technology, internet connectivity, and smart control mechanisms to enable users to operate household appliances using simple voice commands. By utilizing a Wi-Fi- enabled microcontroller such as NodeMCU and a relay module, the system provides seamless communication between the user and connected devices. The implementation of IoT technology allows real-time monitoring and remote control of appliances from any location through a smartphone application. This project demonstrates how automation can simplify daily activities, reduce manual effort, and improve the overall quality of life. It is particularly beneficial for elderly individuals and physically challenged people, as it enables hands-free operation of electrical appliances. The system also contributes to energy conservation by allowing users to manage and monitor appliance usage efficiently, thereby reducing unnecessary power consumption. Furthermore, the project offers flexibility and scalability, making it easy to add more devices and features according to user requirements. The successful operation of lights, fans, and other appliances through voice commands proves the reliability and effectiveness of the proposed system. The integration of cloud services ensures secure data transmission and enhances accessibility. As smart home technologies continue to evolve, IoT-based voice-controlled automation systems are expected to play a significant role in creating intelligent and connected living environments. The project highlights the practical application of emerging technologies in everyday life and demonstrates how automation can improve user experience while promoting energy efficiency and safety. Therefore, Voice Command Based Home Automation Using IoT can be considered a cost- effective, user-friendly, and future-oriented solution that meets the growing demand for smart and automated homes, paving the way for more advanced and intelligent home management systems in the future.

REFERENCES:

[1] Namburi Nireekshana, Tanvi H Nerlekar, P. N. Kumar, and M. M. Bajaber, “An Innovative Solar Based Robotic Floor Cleaner,” May 2023, doi: 10.5281/ZENODO.7918621.
[2] P. K. Balachandran, “Smart IoT in Renewable Energy Systems,” *EEENG*, vol. 17, no. 10, pp. 953–953, Dec. 2024, doi: 10.2174/235209651710240812115711.
[3] A. K. Biswal, D. Singh, B. K. Pattanayak, D. Samanta, and M.-H. Yang, “IoT-Based Smart Alert System for Drowsy Driver Detection,” *Wireless Communications and Mobile Computing*, vol. 2021, no. 1, p. 6627217, Jan. 2021, doi: 10.1155/2021/6627217.
[4] N. Nireekshana, “A POD Modulation Technique Based Transformer less HERIC Topology for PV Grid Tied-Inverter,” in *E3S Web of Conferences*, EDP Sciences, 2025, p. 01001. Accessed: Oct. 29, 2025. [Online]. Available: https://www.e3s-conferences.org/articles/e3sconf/abs/2025/16/e3sconf_icregcsd2025_01001/e3sconf_icregcsd2025_01001.html

[5] N. Nireekshana, R. Ramachandran, and G. V. Narayana, “A Novel Swarm Approach for Regulating Load Frequency in Two-Area Energy Systems,” *IJEER*, vol. 11, no. 2, pp. 371–377, Jun. 2023, doi: 10.37391/ijeer.110218.

[6] P. K. Balachandran, “Smart IoT in Renewable Energy Systems,” *EEENG*, vol. 17, no. 10, pp. 953–953, Dec. 2024, doi: 10.2174/235209651710240812115711.

[7] N. Nireekshana, R. Ramachandran, and G. V. Narayana, “A New Soft Computing Fuzzy Logic Frequency Regulation Scheme for Two Area Hybrid Power Systems,” *International Journal of Electrical and Electronics Research*, vol. 11, no. 3, pp. 705–710, Aug. 2023, doi: 10.37391/IJEER.110310.

[8] N. Nireekshana, R. Ramachandran, and G. V. Narayana, “A Peer Survey on Load Frequency Contol in Isolated Power System with Novel Topologies,” *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 11, no. 1, pp. 82–88, Oct. 2021, doi: 10.35940/ijeat.A3124.1011121.

[9] Namburi Nireekshana, M. Anil Goud, R. Bhavani Shankar, and G. Nitin Sai Chandra, “Solar Powered Multipurpose Agriculture Robot,” May 2023, doi: 10.5281/ZENODO.7940166.

[10] N. Nireekshana, R. Ramachandran, and G. V. Narayana, “Novel Intelligence ANFIS Technique for Two-Area Hybrid Power System’s Load Frequency Regulation,” *E3S Web Conf.*, vol. 472, p. 02005, 2024, doi: 10.1051/e3sconf/202447202005.