RESSTECH

Vol. 03, No. 01, pp. 26-36, 2026

RESEARCH ON INSTRUMENTATION

https://rin.resstech.org/

Design and Implementation of a Hybrid Voice- and IoT-Based Smart Home Control System Using Local Speech Processing to Reduce Cloud Dependency

Sukma Maksum[1*], Yulkifli[1] and Mona Berlian Sari[1]

[1]*Department of Physics, Faculty of Mathematics and Natural Sciences, Padang State University, Padang, Indonesia*

**Corresponding author: sukmamaksum371@gmail.com*

ARTICLE INFO

Article history:
Received May 8, 2026
Revised June 14, 2026
Accepted June 30, 2026
Available online June 30, 2026

Keywords:
Control system; voice recognition; Internet of Things (IoT); Arduino Uno; ESP32; Blynk; smart home.

ABSTRACT

Electrical devices are an important part of modern life. However, forgetting to turn off electrical appliances is still one of the causes of house fires in many areas. This research aims to determine the performance and design specifications of a home electrical appliance control system using a voice recognition sensor based on the Internet of Things (IoT). The research method used is an engineering method with several steps, including problem identification, conceptual design, system design, detailed design, prototype building, and system testing. The system is built using an Arduino Uno microcontroller, ESP32 module, voice recognition sensor, and relay module, which are connected to the Blynk application as a user interface for real-time communication. The system has two modes: voice control mode and manual control mode. In voice control mode, devices work based on voice commands from the user. In manual mode, users can control and monitor devices remotely through the Blynk application using a switch. The test results show that the system achieved an accuracy of 90% and demonstrated good consistency under controlled testing conditions. Furthermore, the system operates effectively and has demonstrated a high level of functional success during testing. These results show that the IoT-based control system using a voice recognition sensor works well and is reliable for controlling home electrical appliances.

1. INTRODUCTION

Rapid transformations in the technological landscape, economic sectors, social structures, and 21st-century activities as a result of increased communication and intelligent automation have led to the term “Industry 4.0” being used to characterize this era [1]. The development of communication technology has attracted significant attention to smart home automation systems. Home automation systems are part of the Internet of Things (IoT), which utilizes Internet technology to connect and manage household devices through smart home automation systems. The system consists of Internet of Things technology, interconnected sensors, and household appliances that can communicate independently with little or no human interaction [2]. Smart home technology refers to devices that provide digitally connected, automated, or enhanced services for building occupants. Smart homes have become a central topic in recent technological and policy discussions regarding energy efficiency, climate change, and building sustainability [3]. The development of smart home technology is a result of technological advancement aimed at helping people manage daily activities in controlling electronic devices and monitoring home conditions. [4].

In Indonesia, household electrical systems are still predominantly operated manually using conventional switches. This condition is considered less efficient because users must interact directly with each device to control its operation. Therefore, the development of more flexible and intelligent control systems is increasingly needed to improve automation in residential environments.

Previous studies have explored smart home systems and voice-based control technologies using various approaches. An IoT-based lighting control system using the NodeMCU ESP8266 integrated with smartphone-based voice recognition has been developed; however, it still depends on the availability of a smartphone and has limitations in the number of controlled devices [5]. IoT-based smart home systems with voice control via a web interface on a smartphone have also been proposed, but their functionality is limited as they rely on a local WiFi network [6]. In addition, a voice-based light control system using the Arduino platform has been developed, but does not yet support internet-based remote control access [7]. The Google Assistant-based system shows good performance, but is highly dependent on a stable internet connection [8]. Recent research proposes an ESP32-based smart home system integrated with Firebase and Android apps with high accuracy, but still relies on external app-based voice processing [9].

In addition to academic research, commercial smart home platforms such as Google Home, Amazon Alexa, and Apple HomeKit have been widely adopted for voice-controlled home automation. These platforms provide advanced speech recognition capabilities, integration with a wide range of IoT devices, and intelligent automation features. However, most commercial voice assistants rely on cloud-based speech processing, where user voice data must be transmitted to remote servers before commands can be interpreted and executed. Although cloud-based processing generally offers high recognition accuracy and natural language understanding, it also introduces concerns related to privacy, security, and user trust. Privacy remains one of the most significant issues associated with smart speaker adoption because voice interactions may be collected, processed, or stored by third-party service providers [10]. Furthermore, concerns regarding security vulnerabilities, data protection, transparency, and trust continue to influence user acceptance of conversational AI technologies and smart voice assistants [11], [12]. Another limitation of cloud-based voice assistant systems is their dependence on internet connectivity. Since voice commands are processed remotely, unstable internet connections may increase response latency or even prevent the system from functioning properly. In addition, commercial ecosystems often require proprietary hardware and dependence on specific service providers, which may increase implementation and maintenance costs [10], [12]. Consequently, despite the rapid advancement of cloud-based smart home technologies, locally processed voice recognition systems remain highly relevant, particularly for low-cost, independent, and privacy-oriented smart home applications.

By processing voice commands directly on embedded hardware, local speech recognition systems can improve privacy protection because voice data does not need to be continuously transmitted to external servers [10]. Local processing also reduces dependence on internet connectivity, enabling the system to continue operating even when network access is limited or unavailable. Moreover, this approach can lower operational costs by minimizing reliance on cloud services while maintaining reliable control of household electronic devices [12]. These advantages make local speech recognition an attractive alternative for developing affordable, efficient, and autonomous smart home systems. Although previous studies have demonstrated the feasibility of voice-controlled smart home systems, most existing solutions either depend on smartphones, local WiFi networks, cloud-based voice processing services, or proprietary commercial ecosystems [5]–[9]. To the best of the authors' knowledge, limited studies have integrated dedicated local voice recognition hardware with IoT-based remote monitoring and control capabilities within a single low-cost smart home platform. This gap indicates the need for a solution that combines the advantages of local speech processing and IoT connectivity while minimizing dependence on external services.

Based on these limitations, this study proposes the development of a smart home control system based on voice recognition integrated with the Internet of Things (IoT) using the Voice Recognition Module V3, Arduino Uno, and ESP32. The system is designed to process voice commands locally via a microcontroller while simultaneously providing remote control and monitoring capabilities through the smartphone-based Blynk application.

The innovation of this research lies in the integration of local speech recognition processing with Internet of Things (IoT)-based remote control in a single system. Unlike cloud-dependent smart home solutions, the proposed system performs voice recognition directly on the embedded device using the Voice Recognition Module V3 while maintaining remote accessibility through IoT technology. This approach provides a more flexible, low-cost, privacy-preserving, and efficient smart home solution by reducing reliance on external voice processing services and increasing system independence in controlling household electronic devices.

2. MATERIAL AND METHODS

The research conducted was engineering research. The procedures involved in the research included developing ideas and task clarity, conceptual design, functional geometric arrangement, detailed design, prototyping, and testing. These stages are commonly applied in IoT-based system development to ensure structured implementation and validation [13][14]. In the research conducted, the data collected included system accuracy, system precision, the effect of smartphone distance on system response time, and system functional testing. Similar studies on IoT-based smart home systems also evaluate performance through accuracy and response time analysis to determine system reliability [15][16]. The system components are integrated according to their respective functions, ensuring smooth and responsive operation, which is consistent with ESP32-based IoT architectures that emphasize modular integration of sensors, microcontrollers, and communication platforms [17]. A block diagram of the home appliance control system using IoT-based voice recognition sensors is shown in Figure 1.

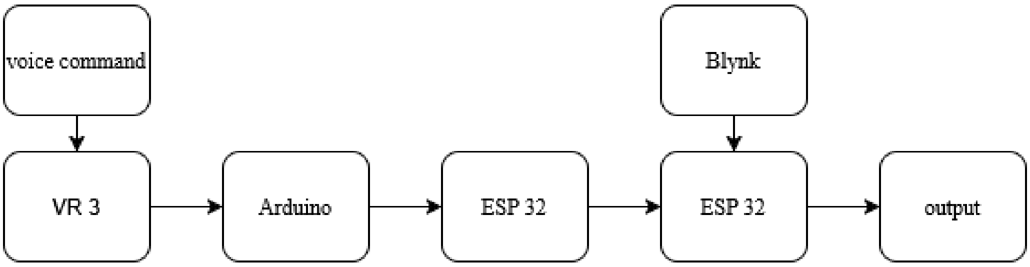

Figure 1. Block diagram of a household electronic device control system using a voice recognition sensor.

In Figure 1, the components used include a voice recognition sensor as a voice input receiver. An Arduino Uno and ESP32 act as microcontrollers to process and control each component, both sensors and actuators. Blynk is used as a mobile-based user interface for monitoring and controlling systems remotely. The system output is an actuator in the form of a relay, which is used as an automatic switch within the system. All components are combined into a single, interconnected system to produce output that meets the system's intended purpose.

Design specifications are important parameters that determine the overall performance and reliability of the system [18]. The hardware system design is divided into two parts, namely the transmitter and receiver, which can be seen in Figure 2 and Figure 3.

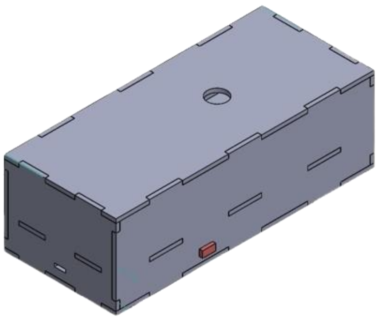

Figure 2. Transmitter

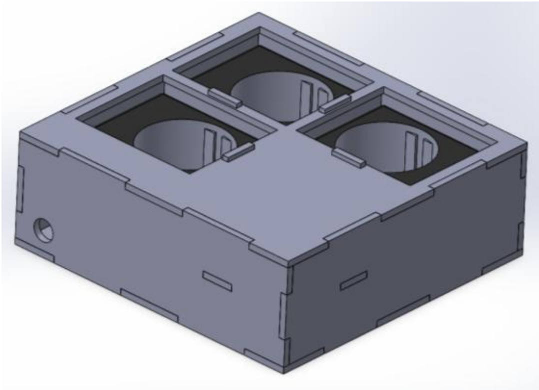

Figure 3. Receiver

The system works with the transmitter as the command receiver, which is then processed by the microcontroller and forwarded to the receiver. The receiver then receives the data and activates the relay according to the command received by the transmitter. The overall system workflow within the microcontroller is displayed in flowchart form. The flowchart shows the data transfer process from input to system operation, as shown in Figure 5.

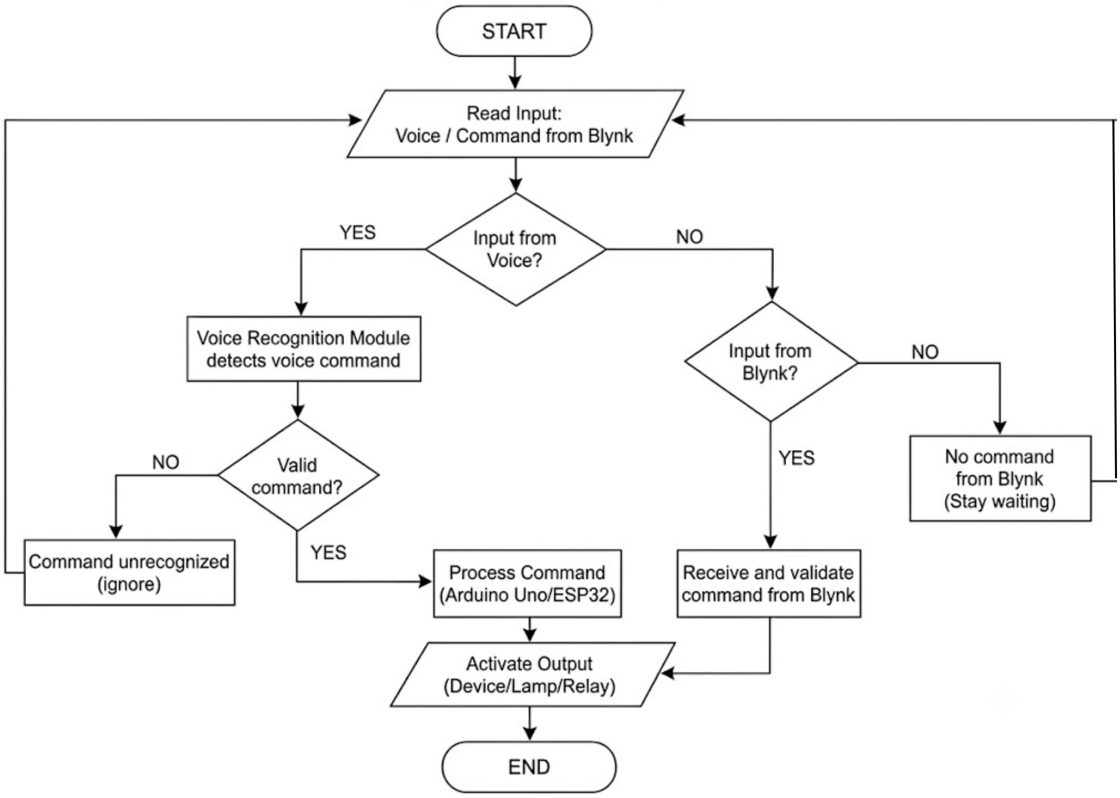

Figure 5. System flowchart

Based on Figure 5, the process starts from reading the voice input that goes into voice recognition or reading the command input from the Blynk application. Then the system will detect the incoming voice to see whether the incoming voice command matches the one in the program. If it matches, the system will process the output according to the voice command, but if not, the system will return to standby to read other voice commands.

Then for the Blynk application, when a command is inputted via the push button of the application, the command will be forwarded and processed by the system according to the input provided. However, if there is no command from the Blynk application, the system will standby to receive commands. To determine the performance and capabilities of the electronic device control system, several tests were conducted, including system accuracy, system precision, and functional testing of the household electronic device control system. This aimed to identify any limitations of the system.

3. RESULTS AND DISCUSSION

The results of this study cover two aspects: performance specifications and system design specifications. Performance specifications detail the shape and dimensions of the hardware, components, and the function of each component used [19]. Based on this, the performance specifications also include the hardware electronic circuit, the components used and the function of each component. The system's electronic circuit is made based on the flowchart shown in Figure 5. Details regarding the system's electronic circuit are presented in Figure 6 for the transmitter and Figure 7 for the receiver. The list of hardware components and materials used in this study is summarized in Table 1 as the Bill of Materials (BOM).

Table 1. Bill of Materials

component	Quantity	Function
Arduino Uno	1	Reading and processing data
ESP 32	2	Controlling the overall system
Voice Recognition Sensor Module	1	Receiving and recording incoming voice commands
Relay	3	As an automatic switch in the system.

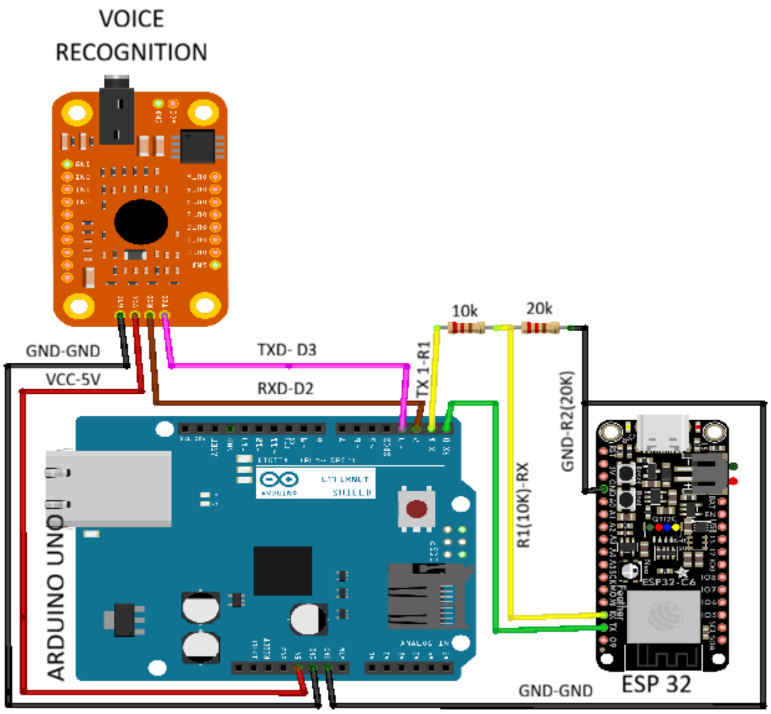

Figure 6. Transmitter electronic circuit

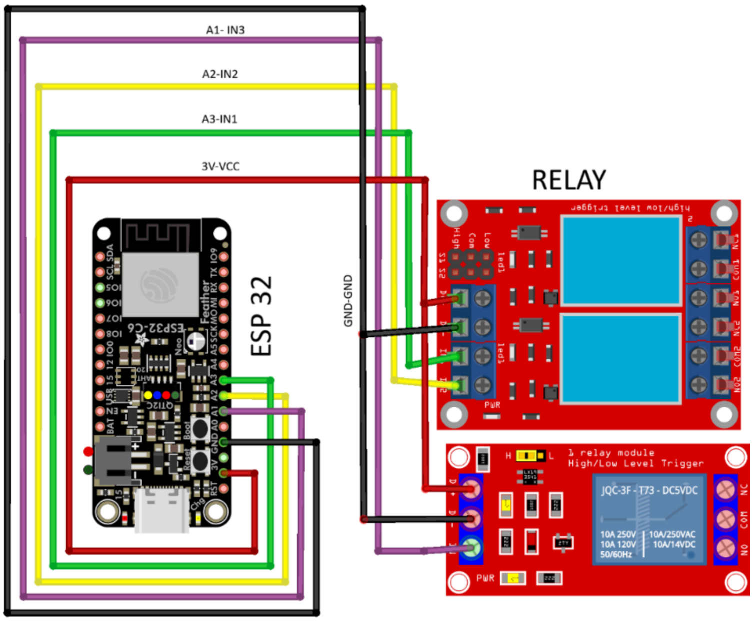

Figure 7. Receiver electronic circuit

Figure 6 shows the electronic transmitter system circuit consisting of a voice recognition sensor, Arduino Uno and ESP32. The voice recognition sensor is used as voice input, Arduino Uno is used to process commands entered into the system, while ESP32 functions as a wireless connection between the system transmitter and receiver. Figure 7 shows an electronic system circuit consisting of an ESP32 and a relay. The ESP32 is used to receive data sent by the transmitter and then processes the relay according to the data sent by the transmitter.

Design specifications are important parameters that reflect the performance capabilities of the equipment based on standard values that must be achieved, as well as the suitability of the equipment's function when testing is carried out [20]. The design specification describes the static characteristics of the product, tolerances, system components, system dimensions, and system dimensions [18]. The system design specifications encompass several aspects, including system accuracy, system precision, and system functionality. The system accuracy test involved 10 trials using different voice commands. The results of the system accuracy test are shown in Table 2.

Table 2. Accuracy of the voice recognition control system

No	Test	Actual orders	System introduction results	Results
1.	1	Turn on	Turn on	Correct
2.	2	Off	Off	Correct
3.	3	*On*	*On*	Correct
4.	4	Close	Close	Correct
5.	5	Open	Open	Correct
6.	6	Turn off	Turn off	Correct
7.	7	Off	Off	Correct

8.	8	Open	Open	Correct
9.	9	Close	Close	Correct
10.	10	Open	On	Wrong

System accuracy is the comparison between the system's recognition results and the actual commands. System accuracy testing is performed by varying the actual commands given to the system and then observing the output of the system's recognition. A system is said to be accurate when the system recognition results match the actual commands given. Based on Table 2, of the ten command variations given, nine system recognitions matched the actual commands. Based on the test results, the system achieved an accuracy of 90%, indicating good overall performance. However, despite this high accuracy, there was a misclassification in one of the trials, where the command "Open" was incorrectly recognized as "On", which may lead to different system outputs. This suggests that although the system performs well in general conditions, there are still limitations in distinguishing phonetically similar commands.

To evaluate the repeatability of the proposed system, the same voice command ON was repeatedly applied under identical testing conditions for ten consecutive trials. This test aimed to determine the consistency of the system's response to repeated inputs. The results of the repeatability test are presented in Table 3.

Table 3. Repeatability Test of the Voice Recognition Control System

No	**Try to**	**Actual conditions**	**Sensor outputs**	**Results**
1	1	*On*	*On*	True
2	2	*On*	*On*	True
3	3	*On*	*On*	True
4	4	*On*	*On*	True
5	5	*On*	*On*	True
6	6	*On*	*On*	True
7	7	*On*	*On*	True
8	8	*On*	*Off*	Wrong
9	9	*On*	*On*	True
10	10	*On*	*On*	True

The test results presented in Table 3 show that the system successfully recognized and executed the repeated voice command in 90% of the trials, with an error rate of 10%. Out of ten repeated trials using the same voice command ("ON"), the system produced the expected output in nine trials, while one trial resulted in an incorrect output ("OFF"). These results indicate that the system was able to provide a relatively consistent response when the same input was repeatedly applied under identical testing conditions. The occurrence of an incorrect response may be attributed to environmental noise, variations in voice pronunciation, or limitations of the Voice Recognition V3 module in distinguishing voice patterns. Nevertheless, the overall results suggest that the proposed system demonstrates good repeatability for voice-based control applications.

The occurrence of an incorrect response may be attributed to several factors, including environmental noise, variations in voice pronunciation, speaking distance, and limitations of the Voice Recognition V3 module in distinguishing similar voice patterns. Such factors can affect the quality of the received voice signal and consequently reduce recognition performance. Despite the presence of one error, the system successfully recognized and executed the repeated voice command in 90% of the trials, indicating a relatively consistent response under controlled testing conditions. The results suggest that the integration of the Voice Recognition V3 module with the ESP32 platform is promising for smart home control applications. However, since the repeatability test was limited to ten trials, the long-term frequency of similar errors could not be determined and should be evaluated through larger-scale testing in future studies. Although only one error was observed, such recognition failures could become critical when controlling safety-sensitive devices such as electric heaters or electric stoves. Therefore, future improvements may include noise reduction techniques, optimized microphone placement, more advanced voice recognition technologies, and additional safety mechanisms such as two-step confirmation, command repetition, or confidence-based verification to reduce false activations and improve system reliability.

To evaluate the effect of smartphone distance on system performance, a response time test was conducted by varying the distance between the smartphone and the smart home system. The response time was measured as the time required for the system to execute a command sent through the Blynk application. The test was performed at several distances ranging from 2 m to 20 m, and the results are presented in Figure 8.

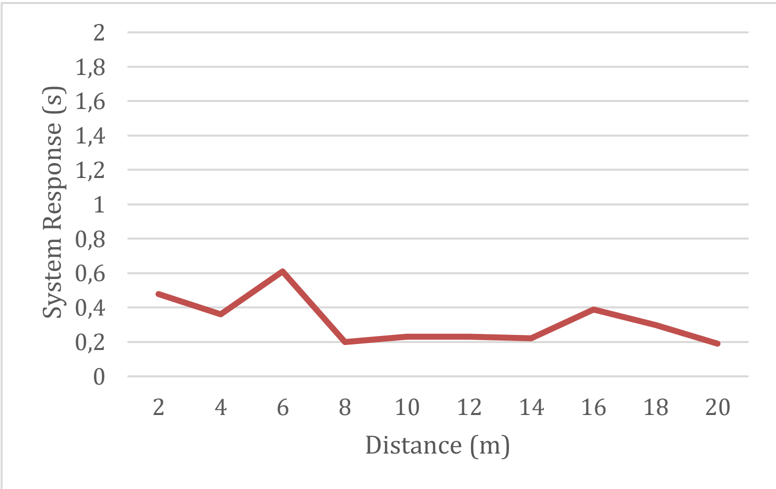

Figure 8. Effect of Smartphone Distance on System Response Time

Based on Figure 8, the system response time ranged from 0.19 s to 0.60 s. The highest response time was recorded at a distance of 6 m (0.60 s), while the lowest response time was observed at a distance of 20 m (0.19 s). Although some fluctuations were observed during the testing process, no consistent increase in response time was found as the distance between the smartphone and the system increased. Overall, the system maintained a response time of less than 1 second across all tested distances. The test results indicate that the system was able to respond to user commands effectively within the tested distance range of 2 m to 20 m. The variations in response time were relatively small and did not exhibit a specific pattern with respect to distance. These findings suggest that the proposed system was capable of operating properly throughout the tested range.

However, this study only evaluated system response time within a limited distance range and did not investigate other factors that may influence system responsiveness. Therefore, future studies should conduct a greater number of trials and evaluate the system under more diverse operating conditions to obtain a more comprehensive understanding of its response-time characteristics.

The next experiment was a functional test of the system. This test aimed to determine the system's ability to perform according to user commands. The functional test observed the correspondence between the voice commands given through the voice recognition sensor and the output response generated by the electronic device's control system. The results of the functional test are shown in Table 4.

Table 4. System functional testing

No	**Voice input**	**Output system**			**Blynk**	**Conclusion**
		Light 1	**Light 2**	**Fan**		
1.	Open	on	off	off	Electronics 1 on	True
2.	Close	off	off	off	Electronics 1 off	True
3.	On	off	on	off	Electronics 2 on	True
4.	Off	off	off	off	Electronics 2 off	True
5.	Turn on	off	off	on	Electronics 3 on	True
6.	Turn off	off	off	off	Electronics 3 off	True
7.	Hidup	off	off	on	Electronics 3 on	True
8.	Mati	off	off	off	*Electronics* 3 off	True
9.	Open	on	off	off	Electronics 1 on	True
10.	On	off	on	off	Electronics 2 on	True

In the system testing process, all voice commands tested were commands that had been previously trained, so they did not involve commands that were not yet recognized by the system. Based on the test results, the system is able to provide an appropriate response to each command given, both to activate and deactivate electronic devices such as Lamp 1, Lamp 2, and Fan. In addition, any changes in the device's condition can be monitored through the Blynk application in real-time, which shows that the system not only has good voice recognition capabilities, but is also integrated with the monitoring system. This indicates that the designed system has worked according to its expected function in controlling electronic devices based on voice commands. Based on the data in Table 3, the control system was successful with an error percentage of 0% and an accuracy percentage of 100%. Based on the data in Table 4, it can be concluded that the results of the functional testing of the electronic device control system using the voice recognition sensor ran well.

The results indicate that the integration of ESP32, Voice Recognition V3, and Blynk provides an effective solution for smart home applications. The system successfully controlled electronic devices through voice commands while simultaneously providing real-time monitoring via the Blynk platform. Although the functional test achieved 100% accuracy, the scalability of the system may be limited by the number of available GPIO pins on the ESP32 and the storage capacity of the Voice Recognition V3 module for voice templates. The Voice Recognition V3 module can store a maximum of 80 trained voice commands, which may become a limiting factor when developing smart home systems with a larger number of devices and functions. As the number of controlled devices increases, additional hardware such as I/O expanders or multiplexers may be required to extend the available control channels. Furthermore, efficient management of voice commands is necessary due to the limited memory capacity of the voice recognition module. Therefore, the proposed system is suitable for small- to medium-scale smart home applications and has the potential to be further developed for larger-scale implementations.

4. CONCLUSION

Based on the research results, a household electronic device control system based on voice recognition and the Internet of Things (IoT) was successfully designed and implemented. The system demonstrated a 90% accuracy rate in recognizing a variety of voice commands, as well as a 90% accuracy rate in repeated tests with the same input. Functional testing showed a 100% success rate, where all devices were controlled according to the given commands and monitored in real-time via the Blynk application.

These results indicate that the system shows good performance under controlled laboratory conditions. However, a misclassification was observed during testing, where the command “Open” was incorrectly recognized as “On”, which may lead to different system outputs. This suggests that the system still has limitations in distinguishing phonetically similar commands and may be affected by factors such as noise interference and variations in speech input during the recognition process.

Despite this limitation, the system was able to perform its intended functions effectively. However, further validation in real residential environments is required before the system can be considered fully reliable. Future studies should conduct long-term testing, for example over a period of at least seven days, with automatic logging of all voice commands and system responses to evaluate reliability under actual operating conditions. Therefore, the proposed system shows potential for further development as an IoT-based smart home solution.

ACKNOWLEDGMENTS

The author would like to thank his supervisor for his guidance and direction throughout this research. He also thanks the examiners for their constructive input and suggestions for improving this research. In addition, the author would like to thank the Physics Study Program and Physics Laboratory for providing facilities and support during the implementation of the research. The author also appreciates all parties who have helped directly or indirectly in completing this research.

DECLARATIONS

Authorship contribution

Sukma Maksum: conceptualization, system design, implementation, testing, data analysis, and script writing

Yulkifli and Mona Berlian Sari: Provide direction, input, and improvements to data and manuscript writing.

Competing Interest

This research does not involve any conflict of interest.

Funding statement

This research did not receive any specific funding from any party.

Ethical Clearance

This research did not involve human or animal subjects.

REFERENCES

[1] Hasan, T., Abrar, M. A., Saimon, M. R., & Sayeduzzaman, M. (2023). Constructing An Integrated IoT-based Smart Home with An Automated Fire and Smoke Security Alert System. Malaysian Journal of Science, 1-10.

[2] Abdulla, A. I., Abdulraheem, A. S., Salih, A. A., M. Sadeeq, M., & Ahmed, A. J. (2020). Internet of Things and Smart Home Security. Technology Reports of Kansai University, 2465-2476.

[3] Sovacool, B. K., & Del Rio, D. F. (2020). Smart home technologies in Europe: A critical review of concepts, benefits,risks and policies. Journal Elsevier.

[4] Bita IM, Hermelingmeier D, Gröger S, Hovemann A, Pfeifer S, Henke C, et al. SmartHomeFarming: Trends, Challenges, and Solutions in a Digital and Sustainable Future. Procedia CIRP. 2025;136:874–9.

[5] Jayusman, Y., Zaenal, R., & Faisal, I. (2020). Perancangan Prototype Kendali Lampu Berbasis Internet of Things (IoT) dengan Nodemcu ESP8266 dan Voice Recognition pada Smarthome. Jurbal Teknologi Informasi dan Komunikasi , 15-25.

[6] Devitra, A., & Purbaningtyas, R. (2022). Prototype Smart Home System Menggunakan Control Pada Perangkat IoT. Jurnal Sistem Informasi, Teknologi Informasi dan Komputer, 53-59.

[7] Cakra, Said, M. S., & Henny. (2023). Sistem Kontrol Lampu Menggunakan Sensor Suara. Jurnal Sistem Informasi dan Teknik Komputer.

[8] Utomo, D., Indriyanto, S., & Yuliantoro, P. (2024). Sistem Otomatisasi Peralatan Rumah Tangga Bagi Penyandang Disabilitas Menggunakan Google Assistant. Jurnal Sistem Informasi dan Teknologi Komputasi, 98-105.

[9] Ramadhana, A., Yulkifli, & Mairizwan. (2025). Development of a Smart Home Control System Based on the Internet of Things Using an Android Application. Reseach On Instrumentation Vol. 2.

[10] Maccario, G., & Naldi, M. (2023). Privacy in Smart Speakers: A Systematic Literature Review. Security and Privacy, 6(1), e274.

[11] Brause, S. R., & Blank, G. (2024). There Are Some Things That I Would Never Ask Alexa: Privacy Work, Contextual Integrity, and Smart Speaker Assistants. Information, Communication & Society, 27(1), 182–197.

[12] Leschanowsky, A., Rech, S., Popp, B., & Bäckström, T. (2024). Evaluating Privacy, Security, and Trust Perceptions in Conversational AI: A Systematic Review. Computers in Human Behavior, 159, 108344.

[13] Jusdi, C. A., Syam, N., & lainnya. (2023). Rancang Bangun Smart Room Menggunakan Voice Recognition Berbasis ESP32 dan Blynk. Journal System Information and Computer, 1(1), 35–46.

[14] Rahayu, P., Sularno, & Sari, I. U. (2023). Perancangan Sistem Smart Home Berbasis IoT Menggunakan ESP32 dan Aplikasi Blynk. Jurnal Sistem Informasi dan Informatika, 3(2), 95–100.

[15] Bhaganagare, S., Chavan, S., Gavali, S., & Godase, V. (2025). Voice-Controlled Home Automation with ESP32: A Systematic Review of IoT-Based Solutions. Journal of Microprocessor and Microcontroller Research.

[16] Fikri, A. (2025). Simple IoT-Based Home Security System Using ESP32 and Blynk. Data Science: Journal of Computing and Applied Informatics, 9(2), 46–50.

[17] Bohara, B., Maharjan, S., & Shrestha, B. R. (2020). IoT Based Smart Home using Blynk Framework. arXiv preprint.

[18] Sarabiti, I., Junaedy, & Martani, A. (2025). Perancangan Prototype Alat Pengukur Energi Listrik Berbasis IoT Menggunakan ESP32 dan Blynk. Jurnal Ilmu Komputer dan Teknologi Informasi.

[19] Mardani, Yohandri, & Kamus, Z. (2016). PEMBUATAN ALAT UKUR DEBIT AIR MENGGUNAKAN SENSOR ALIRAN BERBASIS MIKROKONTROLER ATMEGA328P. PILLAR OF PHYSICS, Vol. 8., 105-112.

[20] Cahyono, B. E., Utami, I. D., Lestari , N. P., & Oktaviany, N. S. (2019). Karakterisasi Sensor LDR dan Aplikasinya pada Alat Ukur Tingkat Kekeruhan Air Berbasis Arduino UNO. Jurnal Teori dan Aplikasi Fisika Vol. 7, No. 2,, 179-186.