

Design and Implementation of an Automatic Door Security System Based on Image and Sound Processing

Najma Hervina Pulungan ^{a,1}, Ojak Abdul Rozak ^{a,2*}, Heri Kusnadi ^{b,3}, Barru Arrosyid ^{b,4}

^a Teknik Elektro, Universitas Pamulang, Jl. Witana Harja No. 18b, Pamulang, Kota Tangerang Selatan, Banten, Indonesia, 15417

^b Teknik Elektro (Kampus Serang), Universitas Pamulang, Jl. Raya Jakarta Km 5 No.6, Kalodran, Kec. Walantaka, Kota Serang, Banten, Indonesia, 42183

e-mail: ¹najmahervinaaaa@gmail.com, ²dosen01314@unpam.ac.id,

³dosen00931@unpam.ac.id, ⁴dosen00859@unpam.ac.id

*Corresponding author

Submitted Date: January 8th, 2026

Reviewed Date : January 26th, 2026

Revised Date: April 27th, 2026

Accepted Date: May 14th, 2026

Abstract

Conventional home security systems relying on mechanical keys are increasingly inadequate and highly vulnerable to unauthorized access. To address these limitations, this study presents the design and implementation of an automated smart door security system that integrates image and voice processing to optimize both security and user convenience. The proposed architecture employs a dual-authentication mechanism: facial recognition via an ESP32-CAM module serves as the primary tier, while voice recognition, facilitated by an Arduino Bluetooth interface and an HC-05 module, functions as the secondary tier. The prototype was empirically evaluated for its accuracy and responsiveness under various environmental conditions. Experimental results indicate that the real-time facial recognition module operates with high precision under standard indoor lighting. While voice authentication exhibited peak accuracy in low-noise environments, its reliability degraded significantly amidst elevated ambient noise. Overall, the integrated electromechanical system demonstrated robust performance in executing automated access protocols. These findings offer a scalable and efficient solution for modern smart home infrastructure, establishing a foundational framework for future advancements in multimodal biometric security systems.

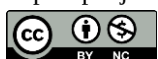
Keywords: Automatic door security; ESP32-CAM; face recognition; voice recognition; Bluetooth Control

1. Introduction

Home door security systems are still predominantly reliant on conventional locking mechanisms, which offer low security levels and are less practical for users [1]. The use of physical keys requires homeowners to carry multiple keys and poses the risk of forgetting or losing them, which can potentially reduce overall security [2]. Moreover, the advancement of door break-in techniques has made it increasingly easier for burglars to unlock conventional doors [3]. These conditions emphasize that home security is a fundamental and essential need [4].

Along with technological advancements, various smart and automatic locking systems have been developed to enhance security and user convenience [5]. Modern security systems are generally designed to adapt and automatically update security mechanisms without requiring manual intervention from users [6]. One widely adopted technology is facial recognition based systems utilizing image processing and computer vision, enabling computers to recognize visual objects like human vision [7]. In addition, voice-based security systems have also been developed to improve flexibility and ease of access [8].

The state of the art indicates that most prior studies have focused on a single authentication method, such as facial or voice recognition, implemented in isolation [9], [10], [11], [12]. Systems that integrate both authentication methods remain limited, particularly in low-cost, microcontroller-based implementations with wireless communication suitable for residential environments [13]. Furthermore,



only a few studies have combined image and voice processing within an automatic door security system based on the ESP32-CAM and Bluetooth communication, which is considered one of the best solutions [14].

Based on these conditions, a research gap exists for a door security system that integrates facial and voice recognition into a single, simple, efficient, and easily implementable system. The integration of dual authentication methods is expected to enhance security levels compared to single-authentication systems while maintaining user convenience.

Therefore, the objective of this study is to design and implement an automatic door security system based on image processing for facial recognition and voice processing for voice command recognition using the ESP32-CAM module and the HC-05 Bluetooth module. This system is expected to enhance security, reduce the risk of unauthorized access, and provide a smarter, more efficient door-locking solution that aligns with societal needs in the digital era [15].

2. Research Methodology

The methodological block diagram used in the design and implementation process of the automatic door security system based on image and voice processing is shown below. This block diagram illustrates the systematic steps of the research process.

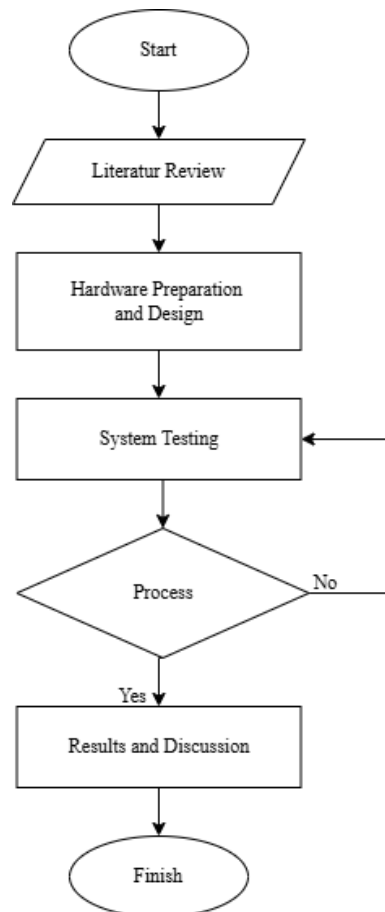


Figure 1. Research Flow Diagram

The research method used in this study consists of several stages, namely: (1) Literature Study: Reviewing previous studies and relevant references related to automatic door security systems, image processing, and voice recognition. (2) System Design: Designing the hardware and software architecture of the system, including the ESP32-CAM, HC-05 Bluetooth module, relay module, and solenoid door lock. (3) Implementation: Realizing the system design in the form of a prototype and performing programming

using the Arduino IDE. (4) Testing: Evaluating system performance based on facial recognition accuracy, voice recognition accuracy, and system response time. (5) Result Analysis: Analyzing the test data to evaluate the reliability and effectiveness of the system.

To implement this system, the design and development of the automatic door security system based on image and voice processing requires several components during the construction process.

A system block diagram illustrating the interaction among several components in a project or system is presented below. The following is an explanation of each element in the diagram.

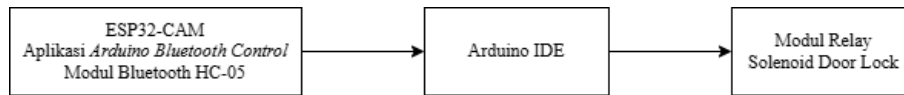


Figure 2. Automatic Door Security System Block Diagram

The block diagram of the automatic door security system based on image and voice processing is shown in Figure 2. This system consists of several main components, namely the ESP32-CAM as the device for detecting and identifying user faces, the HC-05 Bluetooth module for receiving commands from the Arduino Bluetooth Control application, as well as image processing and voice processing units that function to analyze registered faces and recognize user voice commands [16], [17]. The results of these processes are then used to control the door system so that it can automatically open and close.

The circuit of the automatic door security system based on image and voice processing is shown in Figure 3.

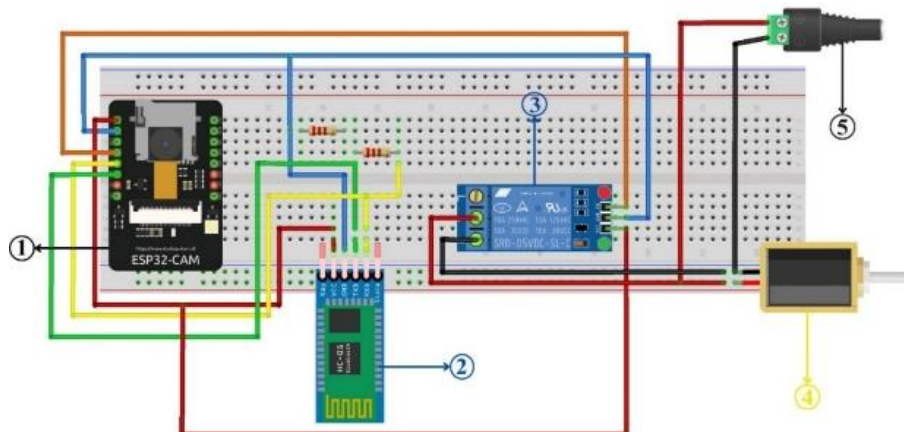


Figure 3. Automatic Door Security System Circuit

Description of Figure 3 (1) ESP32-CAM, (2) HC-05 Bluetooth Module, (3) Relay Module, (4) Solenoid Door Lock, and (5) DC Power Jack (Female).

The figure illustrates a flowchart of the system's operational process, starting from system initialization to data processing and decision making. The process includes image acquisition by the ESP32-CAM camera and reception of voice commands through the HC-05 Bluetooth module connected to the Arduino Bluetooth Control application [18]. This flowchart is designed to facilitate understanding of the overall system workflow, particularly during device initialization, voice command recognition, facial image processing, and door locking control.

The system working diagram of the automatic door security device, based on image and voice processing as follows: (1) The process begins when the system is activated. (2) The system starts detecting objects using the camera. (3) The system analyzes the captured image to recognize faces. (4) The server database stores registered user faces. If a detected face matches the data in the database, the process continues; otherwise, the system returns to the initial state. (5) The Arduino Bluetooth Control application identifies the spoken word. (6) If the word is recognized in the server database, the process continues; otherwise, the system issues a warning and returns to image identification. (7) The system combines the

results of image and voice analysis to determine whether the door should be opened or remain closed. (8) The process ends.

The design of the automatic door security system based on image and voice processing is shown in Figure 4.



Figure 4. Device Design

The design of the automatic door security system based on image and voice processing applied to a room door. In general, the door has a width of 90 cm and a height of 210 cm, representing standard door dimensions. On the right side of the door, there is a control box measuring approximately 12 cm × 15 cm, which functions as a housing for the main system components.

3. Result and Discussion

This chapter discusses the stages of system design, operating principles, testing, and data acquisition of the automatic door security system based on image and voice processing. The system workflow is explained step by step, starting from the detection process to decision making, to improve the overall security level of the system.

a. System Operating Principle

The automatic door security system based on image and voice processing operates by integrating facial identification and voice command authentication as a dual security mechanism. The process begins with capturing the user's facial image using the camera on the ESP32-CAM module, which is then processed using a facial recognition algorithm to match the captured data with the stored database. In addition, the system receives user voice commands through the Arduino Bluetooth Control application, which are transmitted to the Arduino UNO microcontroller via the HC-05 Bluetooth module.

Voice data serves as an additional authentication parameter. The results of facial and voice recognition are then verified by the system to determine user access rights. If both parameters are successfully verified, the system activates the relay module to control the solenoid door lock, allowing the door to open automatically. Conversely, if the verification process fails, access is denied, and the door remains locked as a security measure.

Hardware assembly in this research represents the stage of constructing each part of the device, including the enclosure that houses the system components such as the ESP32-CAM, HC-05 Bluetooth module, relay module, and solenoid door lock.

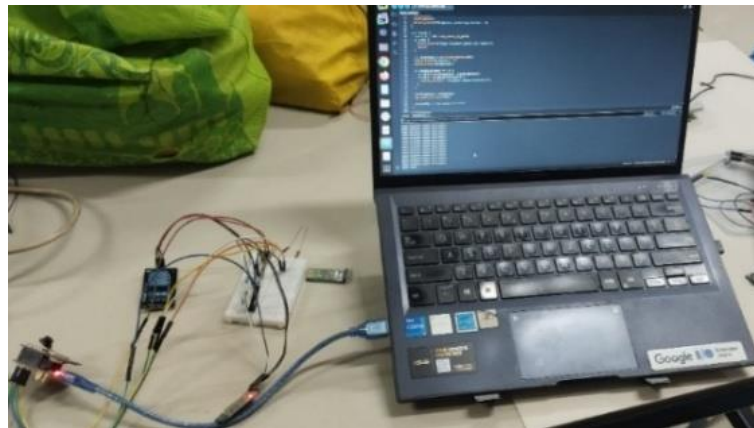


Figure 5. Hardware Assembly

Figure 5 shows the process or result of the hardware assembly of the developed system, namely the automatic door security system based on image and voice processing. In this research, hardware assembly is a crucial stage to integrate all physical components used in the system.

Software development involves designing the program used to process system commands. The software was developed using the Arduino IDE as the main programming platform, where the designed program is uploaded to the ESP32-CAM board, while voice commands are transmitted via the Arduino Bluetooth Control Android application through the HC-05 Bluetooth module.

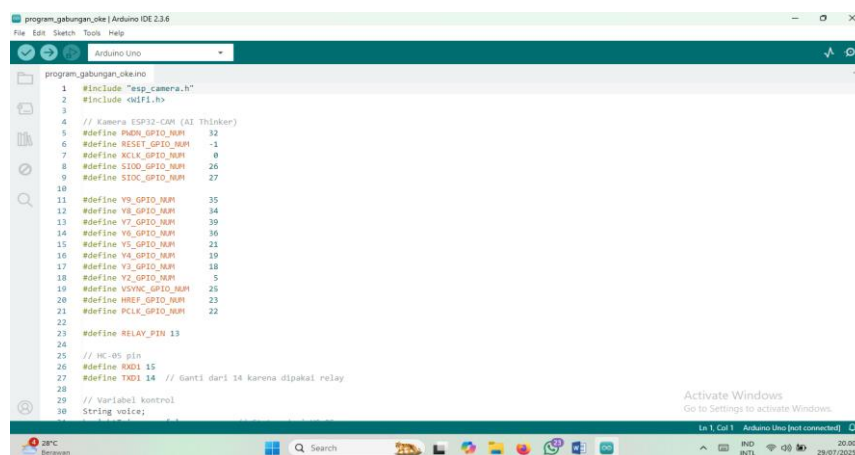


Figure 6. Software Development

Figure 6 shows the interface of software development using the Arduino IDE. This program is part of the system development process for the automatic door security system based on image and voice processing. The display shows the program code written for the ESP32-CAM module, which functions as the facial recognition sensor.

After completing the hardware assembly and software development, system integration is required to ensure that both hardware and software operate as intended. In this research, system integration involves uploading the program developed in the Arduino IDE to the ESP32-CAM board and connecting the Arduino Bluetooth Control application via the HC-05 Bluetooth module. If this process runs smoothly, it indicates that both hardware and software function properly. Hardware testing aims to verify whether the device operates according to the programmed instructions and whether each command is correctly executed by the respective components.

Testing of the ESP32-CAM module was conducted to ensure that the camera functions properly in capturing images and transmitting them to the processing system or user. This module serves as the primary image sensor that detects the presence of a user in front of the door and is integrated with the microcontroller system, as shown in Figures 7 and 8.



Figure 7. ESP32-CAM with Web Server

Figure 7 shows the testing process of the ESP32-CAM module connected to a laptop and operated through a web server. The ESP32-CAM functions as a microcontroller equipped with a camera, enabling it to capture images and transmit them via a WiFi network.

The figure shows the ESP32-CAM module connected to a laptop using a USB cable for programming and power supply.

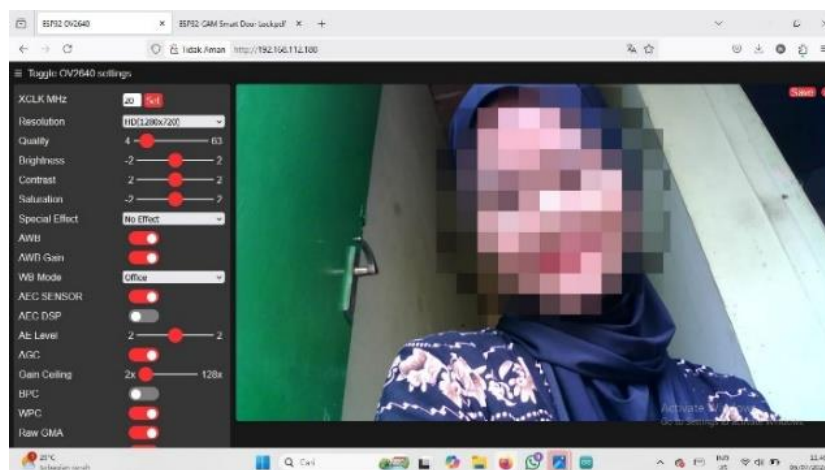


Figure 8. ESP32-CAM Web Server Interface

Figure 8 shows the display on the laptop screen presenting the ESP32-CAM web server interface accessed through a browser. This interface provides various camera configuration parameters, such as image resolution, brightness, contrast, and image quality. In addition, the web server displays real-time camera output (live streaming), allowing users to monitor captured images directly.

Testing of the Arduino Bluetooth Control application was performed to ensure that the system can recognize and process voice commands accurately according to predefined keywords. This module functions as a secondary security sensor that allows the door to unlock only when the spoken command matches the recorded data, as shown in Figure 9.



Figure 9. Arduino Bluetooth Control Application Testing

Figure 9 shows that the Arduino Bluetooth Control application has been tested and can receive commands from a smartphone via a Bluetooth connection. This test confirms that communication between the Android application, Bluetooth module, and ESP32-CAM microcontroller operates properly, indicating that the system is ready for use in electronic device control or automation applications.

Testing of the solenoid door lock was conducted to ensure that the door locking actuator operates optimally according to control signals from the ESP32-CAM and HC-05 Bluetooth module via the relay module. The solenoid door lock functions as the final output component that locks or unlocks the door based on commands from the image and voice processing system, as shown in Figure 10.

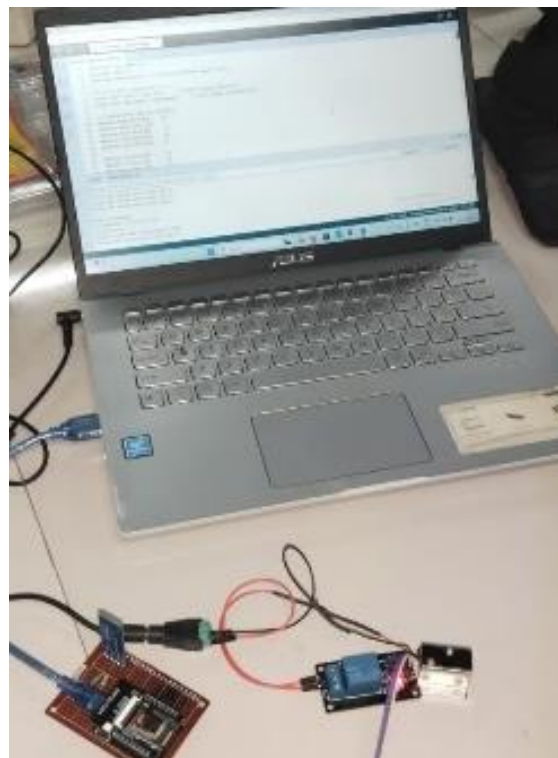


Figure 10. Relay Module and Solenoid Door Lock Testing

Figure 10 shows that the relay module and solenoid door lock were successfully tested and function according to commands from the microcontroller. This test demonstrates that the electronic-based door locking system operates properly and is ready to be integrated into the overall security system.

b. Data Acquisition

Data acquisition was conducted to obtain performance information from each component of the automatic door security system based on image and voice processing. The collected data includes sensor responses (ESP32-CAM and Arduino Bluetooth Control application via the HC-05 Bluetooth module), system response time, and actuator performance (solenoid door lock). These data are used to evaluate whether the system operates effectively, quickly, and accurately under real-world conditions.

The purpose of data acquisition in this chapter is to measure the performance and accuracy of the ESP32-CAM in facial image processing for the automatic door security system. Testing was conducted under various lighting conditions, distances, and system response times during the identification and authentication process.

During this testing, illumination levels were measured in lux, ranging from dark conditions (0 lux) to bright conditions (110 lux).

Table 1. Lighting Effects

Light Conditions (Lux)	Face Detection	Authentication	Notes
110	Detected	Successful	Optimal
105	Detected	Successful	Optimal
93	Detected	Successful	Optimal
88	Detected	Successful	Optimal
80	Detected	Successful	Optimal
77	Detected	Successful	Optimal
65	Detected	Successful	Optimal
42	Undetectable	Fail	Not Optimal
10	Undetectable	Fail	Not Optimal
0	Undetectable	Fail	Not Optimal

Table 1 explains the effect of lighting levels on the system's success in detecting objects using the ESP32-CAM camera sensor. Under low lighting conditions (<42 lux), the system has trouble capturing clear images, resulting in low detection success. In contrast, under moderate to bright lighting conditions (65–110 lux), the system can detect objects with high accuracy. This indicates that system performance is highly influenced by environmental lighting conditions.

Data acquisition was also conducted to determine the success rate, accuracy, and response speed of the Arduino Bluetooth Control application used in the automatic door security system based on image and voice processing. Testing was performed at various voice distances using the application.

To evaluate the effect of environmental noise on voice recognition accuracy, tests were conducted under various noise conditions, ranging from quiet environments to highly noisy environments. The results of these tests are presented in Table 2.

Table 2. The Effect of Background Noise

Decibel (dB) Condition	Sound Detection	Authentication	Notes
44.31	Detected	Successful	Optimal
53.78	Detected	Successful	Optimal
59.79	Undetectable	Fail	Not Optimal
76.18	Undetectable	Fail	Not Optimal
79.68	Undetectable	Fail	Not Optimal
81.81	Undetectable	Fail	Not Optimal
86.47	Undetectable	Fail	Not Optimal

Table 2 explains the effect of environmental noise levels on the success of command recognition by the Arduino Bluetooth Control application. Based on the data, it can be observed that at low noise levels (44.31 dB), recognition accuracy is very high. However, as ambient noise intensity increases (59.79 dB), system accuracy decreases significantly. This demonstrates that the application is more effective in quiet environments or environments with good noise control.

Testing of the solenoid door lock was conducted by repeatedly activating and deactivating the solenoid seven times. The average response time, voltage stability, and mechanical consistency (smoothness of solenoid movement) were measured. The results of this testing are presented in Table 3.

Table 3. Stability Test and Response Time		
Test to-	Response Time (sec)	Status
1	1.3	Successful
2	1.4	Successful
3	1.3	Successful
4	1.2	Successful
5	1.3	Successful
6	1.1	Successful
7	1.2	Successful

Based on Table 3, it can be observed that the response time does not change significantly despite repeated testing. This indicates that the solenoid door lock has good mechanical and electrical stability and is capable of maintaining consistent performance. Based on these results, it can be concluded that the solenoid door lock is suitable for systems requiring fast response times and stable long-term operation.

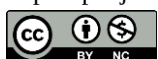
4. Conclusion

Based on the information and data obtained during the undergraduate thesis research on an automatic door security system based on image and voice processing, the following conclusions can be drawn (1) The system is capable of detecting user faces in real time using the internal camera of the ESP32-CAM. Test results indicate that under moderate lighting conditions (above 65 lux), the facial detection success rate reaches an optimal level. (2) Valid voice commands activate the second authentication mechanism, in which signals from the Arduino Bluetooth Control application are transmitted through the HC-05 Bluetooth module and processed via the Arduino IDE to activate the actuator. (3) Testing results show that the highest voice recognition accuracy is achieved in a quiet environment (44.31 dB), while the accuracy decreases significantly as environmental noise intensity increases (59.79 dB).

Building upon the findings of the present study, several avenues for future research are proposed to further optimize the system's performance, security, and reliability. First, to fortify the communication link between the smartphone and the microcontroller via the HC-05 Bluetooth module, subsequent iterations should implement advanced cryptographic protocols to mitigate vulnerabilities associated with signal interception and unauthorized access. Additionally, upgrading the hardware infrastructure with high-resolution optical cameras and high-fidelity microphones is recommended to ensure consistent detection accuracy under adverse environmental conditions, including extended spatial ranges, low-light scenarios, and reduced vocal intensity. Finally, future architectures should incorporate real-time diagnostic feedback mechanisms that notify users of specific authentication failures, enabling them to efficiently identify and rectify the underlying causes of improper system responses.

References

- [1] J. Jhonatan and K. Sekarsari, "Sistem Keamanan Dua Lapis Dengan RFID dan Pendeteksi Objek Dengan Machine Learning," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 2, pp. 1059–1070, Sep. 2024, doi: <https://doi.org/10.47065/bits.v6i2.5747>.
- [2] A. Dwi Wibowo and S. Susanto, "Perancangan Purwarupa Smart Door Lock berbasis Internet of Things," *Jurnal Elektro & Informatika Swadharma (JEIS)*, vol. 04, no. 02, pp. 1–8, 2024.



- [3] Mutiara Syafrizal, Vidya Ikawati, Agus Siswanto, and N. Lestari, "Sistem Keamanan Door Lock berbasis Voice Recognition dengan Natural Language Processing," *Infotronik : Jurnal Teknologi Informasi dan Elektronika*, vol. 9, no. 1, pp. 1–11, Jun. 2024, doi: <https://doi.org/10.32897/infotronik.2024.9.1.3611>.
- [4] Wisnu Wiraditama, Dedi Candra, Ami Prasena Nugraha, and Sulartopo Sulartopo, "Sistem Keamanan Rumah Menggunakan Sensor Passive Infrared," *Seminar Nasional Teknologi dan Multidisiplin Ilmu (SEMNASTEKMU)*, vol. 3, no. 1, pp. 200–210, Oct. 2023, doi: <https://doi.org/10.51903/semnastekmu.v3i1.215>.
- [5] I. Supriadi, A. Mauluddin, and H. R. Warits, "Pengembangan Aplikasi Manajemen Keamanan Lingkungan RW 11 Kelurahan Sekejati Kota Bandung Berbasis Android," *Jurnal Tekno Insentif*, vol. 19, no. 1, pp. 107–121, Jul. 2025, doi: <https://doi.org/10.36787/jti.v19i1.1943>.
- [6] Y. Risayani, S. Japit, C. Bombongan, T. Selamat, and Y. Yuliana, "Sistem Keamanan Siber Adaptif Berbasis AI: Analisis Kinerja, Arsitektur, dan Penerapannya pada Organisasi Modern," *Jurnal Minfo Polgan*, vol. 14, no. 2, pp. 2999–3006, Nov. 2025, doi: <https://doi.org/10.33395/jmp.v14i2.15630>.
- [7] Supiyandi Supiyandi, Tegar Ardiansyah, Sri Putri Balqis, Jundi Haqqoni, and Salsa Nabila Iskandar, "Deteksi Wajah dalam Foto Menggunakan Teknologi Visi Komputer," *Mars : Jurnal Teknik Mesin, Industri, Elektro Dan Ilmu Komputer*, vol. 2, no. 6, pp. 32–39, Nov. 2024, doi: <https://doi.org/10.61132/mars.v2i6.490>.
- [8] K. Suy, Sufajar Butsianto, and Supriyanto, "Sistem Informasi Keamanan Kunci Pintu Pintar Berbasis Internet of Things Notifikasi Real-Time Berbasis WhatsApp," *Bulletin of Computer Science Research*, vol. 5, no. 5, pp. 886–897, Aug. 2025, doi: <https://doi.org/10.47065/bulletincsr.v5i5.720>.
- [9] Y. Yao, "Research on facial recognition system based on deep learning," in *Applied and Computational Engineering*, EWA Publishing, Feb. 2024, pp. 232–236. doi: <https://doi.org/10.54254/2755-2721/34/20230332>.
- [10] S. Setiyo Prihatmoko, Sumaryanto, and Jarot Dian Susatyono, "Konsep Desain Multimedia Sinestetik sebagai Pendekatan Baru dalam Penciptaan Pengalaman Visual–Audio Imersif," *Jurnal Publikasi Ilmu Komputer dan Multimedia*, vol. 5, no. 1, pp. 197–213, Jan. 2026, doi: <https://doi.org/10.55606/jupikom.v5i1.6135>.
- [11] K. Aizat, O. Mohamed, M. Orken, A. Ainur, and B. Zhumazhanov, "Identification and authentication of user voice using DNN features and i-vector," *Cogent Eng.*, vol. 7, no. 1, pp. 1–21, Jan. 2020, doi: <https://doi.org/10.1080/23311916.2020.1751557>.
- [12] A. Labinjan, S. Baressi Šegota, I. Lorencin, and N. Tanković, "Face Recognition System Using CLIP and FAISS for Scalable and Real-Time Identification," *Mathematical and Computational Applications*, vol. 31, no. 36, pp. 1–24, Mar. 2026, doi: <https://doi.org/10.3390/mca31020036>.
- [13] S. Uppuluri and G. Lakshmeeswari, "Secure user authentication and key agreement scheme for IoT device access control based smart home communications," *Wireless Networks*, vol. 29, no. 3, pp. 1333–1354, Apr. 2023, doi: <https://doi.org/10.1007/s11276-022-03197-1>.
- [14] I. Fitri, C. Dewi, and J. A. Munif, "Otomatisasi Pintu Dengan Menggunakan ESP32-CAM Berbasis Telegram," in *Seminar Nasional Teknologi Informasi dan Bisnis (SENATIB) 2025*, 2005, pp. 823–834.
- [15] E. Boru Zega, W. Ginting, and R. Damanik, "Sistem Keamanan Login menggunakan Metode Autentikasi Dua Faktor dengan Verifikasi Gambar," *International Multidiciplinary Jpurnal*, vol. 1, no. 02, pp. 25–35, 2025.
- [16] K. Z. Yonatan, H. S. Utama, and Y. Calvinus, "Aplikasi Kendali ESP32 Menggunakan Sensor Sidik Jari pada Ponsel sebagai Lapisan Keamanan," *Aviation Electronics, Information Technology, Telecommunications, Electricals, Controls (AVITEC)*, vol. 6, no. 1, pp. 71–78, Feb. 2024, doi: <https://doi.org/10.28989/avitec.v6i1.2049>.
- [17] Ardiansyah. M, Aldi Febryan, Adriani, and Rahmania, "Rancang Bangun Sistem Keamanan Rumah berbasis Telegram menggunakan ESP 32 Cam," *VERTEX ELEKTO: Jurnal Teknik Elektro UNISMUH*, vol. 15, no. 1, pp. 64–71, 2023.
- [18] H. Hardi, "Perancangan Prototype Mobil Remote Control dengan Smartphone Android menggunakan Bluetooth HC-05 berbasiskan Arduino Uno," *Jurnal Pengelolaan Laboratorium Sains dan Teknologi*, vol. 2, no. 1, pp. 13–22, 2022, doi: <https://doi.org/10.33369/labsaintek.v2i1.22003>.

