Journal of Global Engineering Research & Science (J-GERS)
Vol. 05, No. 01, June 2026, pp. 37-44
ISSN: 2830-5906, *e*-ISSN: 2830-5418, DOI: 10.56904/j-gers.v5i1.201

IoT-Based Locker Lock Security Using YOLO Facial Recognition

Legenda Prameswono Pratama[1], Sinka Wilyanti[1], Naufal Eka Putra[1*]
[1*] Department of Electrical Engineering, Faculty of Engineering & Computer Science, Jakarta Global University, Indonesia, 16412

Article Info

Article history:

Received May 8, 2026
Revised June 26, 2026
Accepted June 29, 2026

Keywords:

Facial Recognition
YOLO
ESP32-CAM
Smart Locker
Internet of Things

ABSTRACT

The increasing demand for secure and intelligent storage systems has encouraged the development of biometric-based authentication technologies. This study presents the design and implementation of a smart locker security system based on facial recognition using the YOLO (You Only Look Once) algorithm integrated with an ESP32-CAM module and IoT communication. The system was developed to provide automated user authentication and remote locker control through real-time face detection and recognition. Performance evaluation was conducted by analyzing face detection accuracy, authentication response time, and network communication quality under different distances and lighting conditions. Experimental results showed that the YOLO-based system achieved a a face detection success rate of up to 92% under both bright-light and low-light environments at distances of 30 cm and 100 cm. The detection error ranged from 2.60% to 4.30%, indicating stable and reliable detection performance. Authentication testing revealed that the system performed optimally at 30 cm, with response times ranging from 1.47 s to 2.56 s and error rates below 1.2%. At 100 cm, the response time increased to approximately 9–10 s, accompanied by higher error rates due to reduced facial feature visibility. In addition, network performance evaluation based on latency, packet loss, and throughput demonstrated reliable Wi-Fi communication between the ESP32-CAM and the server, ensuring smooth data transmission during authentication and locker control operations. The results confirm that the YOLO algorithm is effective for real-time facial recognition applications and can be successfully implemented in smart locker security systems. The proposed system provides accurate authentication, reliable communication, and enhanced security, making it suitable for practical access-control applications.

Corresponding Author:

Naufal Eka Putra
Department of Electrical Engineering, Faculty of Engineering & Computer Science, Jakarta Global University, Indonesia, 16412
Email: naufalekaputra03@gmail.com

1. INTRODUCTION

The rapid advancement of information technology and the Internet of Things (IoT) has significantly increased the demand for secure and efficient authentication systems, particularly for locker security in public spaces [1], [2]. Conventional mechanical locks and numeric combination systems have several limitations, including the risk of key loss, duplication, and unauthorized access [3]. Therefore, the development of modern, digital-based locker security systems is crucial to ensure both safety and convenience. The IoT paradigm enables interconnected devices to communicate and exchange information through network infrastructures, allowing real-time monitoring and automated control of security systems. By integrating sensors, embedded controllers, and wireless communication technologies, IoT-based security solutions can provide enhanced accessibility and operational efficiency. Furthermore, the adoption of intelligent authentication mechanisms can reduce human intervention while improving the reliability and responsiveness of access control systems in various environments such as schools, offices, and public facilities.

Recent studies have explored various technologies for locker security, such as Radio Frequency Identification (RFID) [4], Personal Identification Numbers (PIN) [5], and biometric-based systems [6]. Among these, facial recognition offers a non-contact, practical, and user-friendly approach for authentication [7]. Deep learning methods, particularly the You Only Look Once (YOLO) algorithm, have demonstrated high accuracy and fast face detection capabilities [8], [9]. However, most implementations rely on high-performance devices, making their integration into resource-constrained IoT platforms a significant challenge. YOLO performs object detection through single-stage architecture, enabling real-time processing while maintaining robust detection performance [10], [11]. This characteristic makes YOLO particularly suitable for smart security applications that require rapid authentication and low latency [9]. Several studies have successfully applied YOLO-based facial recognition in access control systems, demonstrating promising results in terms of detection accuracy and response time. However, most implementations rely on high-performance devices, making their integration into resource-constrained IoT platforms a significant challenge. In addition, limited research has investigated the practical deployment of YOLO-based facial recognition for locker security applications using low-cost embedded hardware and wireless communication technologies.

Despite the potential of YOLO for facial recognition, there is a research gap in applying this algorithm to locker security systems using IoT devices such as the ESP32-CAM module. While previous studies have demonstrated the effectiveness of YOLO-based facial recognition in smart door lock and access control applications [12], [13], relatively few investigations have focused on locker security environments that require compact, low-cost, and energy-efficient hardware. Limited computational resources and network constraints in IoT environments require careful design and optimization to ensure reliable performance [14]. In addition, maintaining a balance between detection accuracy, authentication speed, and communication reliability remains a significant challenge when deploying deep learning models on embedded systems. Therefore, there is a need for an integrated solution that combines real-time facial recognition with IoT-based monitoring and control while remaining suitable for resource-constrained devices. Addressing this gap, this study focuses on designing and implementing a locker security system that utilizes YOLO-based facial recognition on an ESP32-CAM platform.

The objectives of this research are threefold: to design and develop a facial recognition-based locker security system using YOLO and ESP32-CAM, to evaluate the system's performance in terms of detection accuracy and authentication response time, and to analyze the quality of network communication based on Quality of Service (QoS) parameters. The proposed system integrates computer vision, wireless communication, and embedded hardware to provide an intelligent authentication mechanism for locker access. Unlike conventional locker security systems that rely on physical keys, passwords, or access cards, the proposed approach enables contactless authentication through facial recognition, thereby enhancing both security and user convenience. Furthermore, the evaluation of QoS parameters provides insights into the communication performance of the IoT infrastructure, ensuring that the system can operate reliably in practical deployment scenarios. The proposed system aims to provide a practical, secure, and efficient solution for locker security while contributing to the advancement of computer vision and IoT integration for real-world smart security applications.

2. METHOD

2.1 Research Design

This study employs a design and implementation approach to develop a locker security system based on facial recognition using the YOLO algorithm integrated with the ESP32-CAM module. The research focuses on both hardware and software development, as well as system performance evaluation in terms of face detection accuracy, authentication response time, and network communication quality. Figure 1 illustrates a 3D design of the locker security system developed to improve cabinet protection through the integration of electronic and biometric technologies. The system is installed on the front side of the locker and consists of several interconnected components, including an ESP32-CAM (used as the main microcontroller without Arduino Uno), ESP32-CAM, fingerprint sensor, keypad keyboard, OLED LCD display, relay module, LED indicator, buzzer, and antenna module. The ESP32-CAM (used as the main microcontroller without Arduino Uno) serves as the main controller responsible for processing data and coordinating communication between all connected devices. The keypad is used for password entry, while the fingerprint sensor provides biometric authentication to ensure that only authorized users can access the locker. The OLED LCD display functions as an interface to show system status, authentication results, and user instructions. Meanwhile, the ESP32-CAM module enables wireless communication and camera-based facial recognition features, supported by the antenna to ensure stable signal transmission. The ESP32-CAM is a compact and cost-effective microcontroller module designed for IoT-based image acquisition and monitoring applications [15]. In this study, the module captures facial images in real time and transmits them through a wireless network for YOLO-based facial recognition and authentication. Its low power consumption, compact size, and built-in networking capabilities make the ESP32-CAM an effective platform for implementing smart security systems in resource-constrained environments [16].

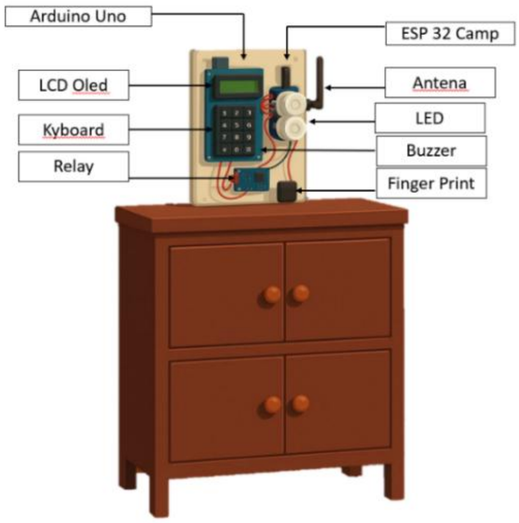


Figure 1. 3D Design of the locker security system

In addition, the LED indicator and buzzer provide visual and audio notifications for successful or failed authentication attempts. The relay module operates as an electronic switch that controls the locking and unlocking mechanism of the locker door. Overall, the 3D illustration demonstrates the structural arrangement and functional integration of the hardware components in developing an intelligent and secure locker security system.

2.2 Hardware and Software Components

The primary hardware used in this study includes the ESP32-CAM microcontroller, a solenoid door lock for locker locking, and supporting power supply and network modules. The software environment consists of Python for YOLO model training and OpenCV for image processing, while the ESP32-CAM firmware is programmed using the Arduino IDE. The YOLOv5 model is selected for its balance between detection speed and accuracy, suitable for deployment on low-resource IoT devices.

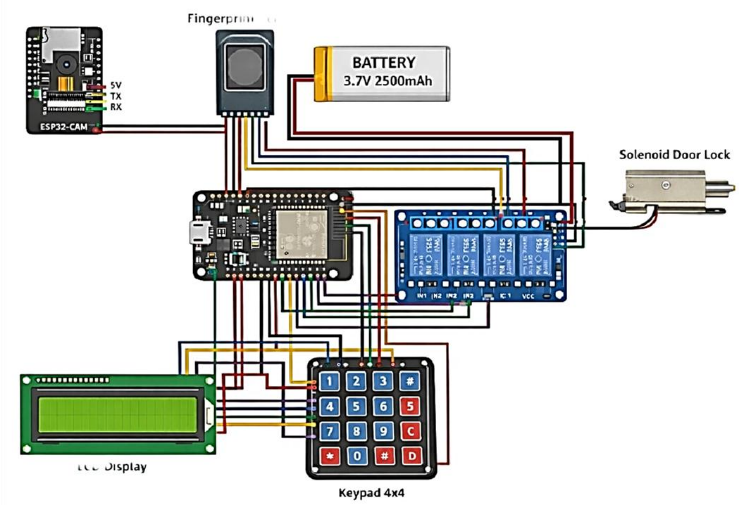


Figure 2. Interconnection of each component

Figure 2 illustrates the interconnection of each component used in the smart door lock system based on the ESP32-CAM microcontroller. The ESP32-CAM functions as the main controller that manages communication and processing between all connected devices. A fingerprint sensor is connected to the ESP32-CAM to provide biometric authentication for user access. A 4×4 keypad is used for password input, while the LCD display provides visual information such as system status and user instructions. The relay module acts as an electronic switch that controls the solenoid door lock, allowing the door to lock or unlock based on successful authentication. All components are powered by a 3.7V 2500mAh battery, ensuring portable and stable operation of the system. The wiring diagram demonstrates how data, control, and power connections are integrated to create a complete electronic door security system.

2.3 System Architecture and Procedure

The system architecture consists of three main components: image acquisition, face recognition, and locker control. Images are captured by the ESP32-CAM and transmitted to the local processing unit or cloud server running the YOLO model. The facial recognition process identifies registered users in real-time. Upon successful authentication, a command is sent to the solenoid door lock to unlock the locker. The system flow is presented in Figure 3.

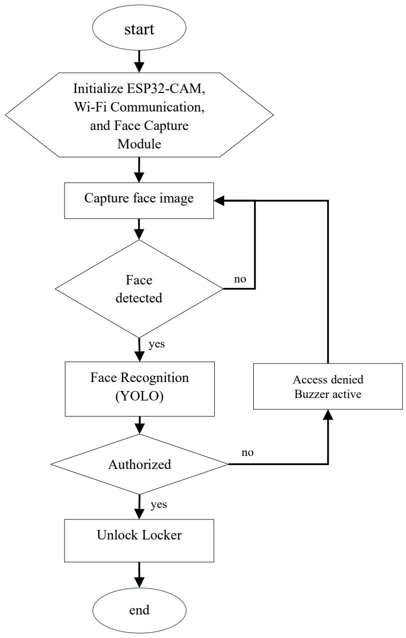


Figure 3. Flowchart of the Facial Recognition Authentication Process

2.4 Data Collection and Training

The facial dataset consists of images of registered users captured under normal indoor lighting conditions. Data augmentation techniques, including rotation, scaling, and brightness adjustment, are applied to increase model robustness. The YOLO model is trained on this dataset, and performance metrics such as precision, recall, and F1-score are recorded.

2.5 Performance Evaluation

System performance is evaluated in three aspects:

1. Face Detection Accuracy: The proportion of correctly identified users compared to total attempts.
2. Authentication Response Time: The time interval between image capture and locker unlocking.
3. Network Communication Quality: Measured using Quality of Service (QoS) parameters, including latency, packet loss, and throughput.

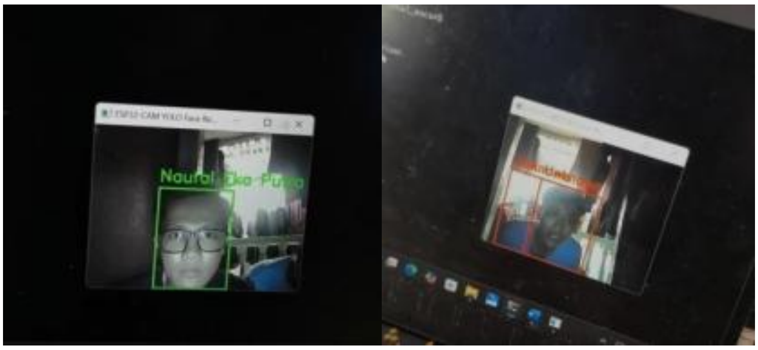
2.6 Limitations of the Method

The system is designed for single-user authentication per session and does not support facial recognition with obstructions such as masks or sunglasses. Only one network connection (Wi-Fi) is utilized, and all model computations are limited by the processing capacity of the ESP32-CAM.

3. RESULTS AND DISCUSSION

3.1. Face Detection Accuracy

Figure 3 presents the performance of the YOLO-based facial recognition system under two different conditions. Figure 3(a) shows a successful face detection and identification process, where the system accurately locates the user's face and displays the correct identity label. The green bounding box indicates that the detected face matches the facial data stored in the database, demonstrating that the YOLO algorithm can effectively detect facial features and perform recognition in real time. This result confirms that the system can authenticate authorized users when facial images are captured under favorable conditions



(a) (b)

Figure 3. Face detection in YOLO algorithm (a) correct, (b) incorrect

In contrast, Figure 3(b) illustrates an unsuccessful recognition result. Although the system can detect the presence of a face, the identification process fails, as indicated by the red bounding box and incorrect labelling. This error may be caused by several factors, including insufficient lighting, variations in facial pose, partial occlusion, image blur, or limitations in the training dataset. Such conditions can reduce the quality of extracted facial features, leading to lower recognition accuracy[17]. The comparison between Figures 3(a) and 3(b) highlights the importance of image quality and environmental conditions in facial recognition systems. While YOLO demonstrates strong capability in face detection, the overall recognition performance depends on the consistency of facial appearance and the robustness of the trained model. Therefore, improving dataset diversity, optimizing illumination conditions, and applying additional preprocessing techniques can enhance the reliability and accuracy of the facial authentication system.

Tabel 1. Face Detection Performance

Metric	Value
Total Image Tested	100
Correctly Detected Faces	92
Incorrectly Detected Faces	8
Accuracy (%)	92

Table 1 presents the performance results of the face detection system based on the conducted testing. A total of 100 images were tested to evaluate the accuracy of the system in detecting faces. The results show that 92 images were correctly detected, while 8 images were incorrectly detected. This indicates that the system achieved an accuracy rate of 92%, demonstrating that the face detection method used in the system performs effectively and reliably under the test conditions. The small number of incorrect detections may have been caused by factors such as poor lighting conditions, variations in facial position, image quality, or partial obstruction of the face during the detection process. Overall, the obtained accuracy confirms that the implemented face detection system can provide satisfactory performance for practical security and authentication applications.

Tabel 2. face detection in distance dan light

Person	Distance (cm)	Condition	Detection Status	Error (%)
1	30	Bright lighting	Yes	2.98
2	100	Bright lighting	Yes	4.10
3	30	Bright lighting	Yes	3.87
4	100	Bright lighting	Yes	3.75
5	30	Bright lighting	Yes	3.84
6	100	Bright lighting	Yes	4.00
7	30	Low light	Yes	2.60
8	100	Low light	Yes	4.20

9	30	Low light	Yes	2.80
10	100	Low light	Yes	4.30

Table 2 presents the performance of the YOLO-based face detection system under different distances and lighting conditions. The experiments were conducted at distances of 30 cm and 100 cm under both bright lighting and low-light conditions. The results show that the system successfully detected all faces, achieving a 100% detection rate across all test scenarios. This indicates that the YOLO algorithm is capable of reliably identifying facial regions regardless of moderate variations in distance and illumination [18]. Under bright lighting conditions, the detection error ranged from 2.98% to 4.10%. The lowest error was observed at 30 cm (Person 1), while slightly higher errors were recorded at 100 cm. This trend suggests that increasing the distance between the camera and the subject may reduce the amount of facial detail captured, leading to a minor decrease in detection precision. Nevertheless, the error values remained below 5%, demonstrating stable performance. Similarly, under low-light conditions, all faces were successfully detected. The detection error varied between 2.60% and 4.30%. Although low illumination generally poses challenges for computer vision systems due to reduced image contrast and increased noise, the YOLO model maintained a high level of detection capability. The slight increase in error at 100 cm compared to 30 cm indicates that the combined effects of distance and limited lighting can influence the quality of detected facial features. The results demonstrate the robustness of the YOLO algorithm for real-time face detection applications. YOLO processes images through a single-stage detection framework, allowing it to simultaneously localize and classify objects with high speed and accuracy [19]. Its ability to detect faces consistently under varying environmental conditions makes it suitable for security and access-control systems, such as the proposed smart locker authentication system. However, the observed increase in error at greater distances and under low-light conditions suggests that detection accuracy can be further improved through enhanced lighting arrangements, higher-resolution image acquisition, and additional training data representing diverse illumination environments. Overall, the findings confirm that YOLO provides reliable face detection performance, achieving consistent detection results while maintaining low error rates in both bright and low-light conditions. This capability supports its implementation as a core component in biometric security systems requiring fast and accurate user authentication.

3.2. Authentication Response Time

Table 3. Authentication Response Time

Person	Distance (cm)	Condition	Delay (s)	Error (%)
1	30	Bright lighting	2.56	1.09
2	100	Bright lighting	9.18	7.71
3	30	Bright lighting	2.16	1.14
4	100	Bright lighting	9.71	8.24
5	30	Bright lighting	1.47	0.00
6	100	Bright lighting	10.39	8.92
7	30	Low light	2.35	0.88
8	100	Low light	9.80	8.33
9	30	Low light	1.68	0.21
10	100	Low light	10.23	8.76

Table 3 presents the authentication response time of the proposed facial recognition system under different distances and lighting conditions. The tests were conducted at distances of 30 cm and 100 cm under both bright lighting and low-light environments. The results indicate that both distance and lighting conditions influence the system's authentication speed, although distance appears to have a greater impact on performance. At 30 cm under bright lighting conditions, the authentication delay ranged from 1.47 s to 2.56 s, with error values between 0.00% and 1.14%. These results demonstrate that the system can authenticate users rapidly when facial features are clearly captured by the camera. The shortest response time of 1.47 s was recorded for Person 5, indicating efficient face detection and recognition under favorable conditions. When the distance increased to 100 cm under bright lighting conditions, the authentication delay increased significantly, ranging from 9.18 s to 10.39 s. The corresponding errors also rose to between 7.71% and 8.92%. This increase can be attributed to the reduced facial detail available at greater distances, requiring additional processing by the YOLO-based detection and recognition system to accurately identify the user. Despite the longer response times, the system remained capable of completing the authentication process successfully. A similar trend was observed under low-light conditions. At 30 cm, authentication delays remained relatively low, ranging from 1.68 s to 2.35 s, with error values below 1%. These findings indicate that the system can maintain stable performance even when illumination is limited, provided that the user's face is sufficiently close to the camera. However, at 100 cm, the response time increased to between 9.80 s and 10.23 s, accompanied by error values exceeding 8%. The combination of greater distance and reduced illumination likely affected image quality, making facial feature extraction more challenging. The observed performance is closely related to the characteristics of the YOLO algorithm. YOLO performs object detection in real time by processing the entire image in a single forward pass through the neural network. Under favorable conditions, such as close distances and adequate lighting, YOLO can quickly detect facial regions and support efficient authentication. However, when facial images become smaller or less distinct due to increased distance or poor lighting, the algorithm requires additional effort to identify facial features accurately, leading to increased processing delays and higher error rates. Overall, the results demonstrate that the proposed system achieves its best authentication performance at 30 cm under both bright and low-light conditions, with response times below 3 seconds and minimal error rates. Although authentication remains successful at 100 cm, the increased delay suggests that closer user positioning is preferable for practical applications. These findings confirm that the YOLO-based facial recognition system is effective for access-control applications, providing reliable authentication performance while maintaining robustness under varying environmental conditions.

3.3. Network Communication Quality

The system uses Wi-Fi for data transmission between the ESP32-CAM and the server. The network quality was evaluated using three parameters: latency, packet loss, and throughput, following standard IoT QoS measurements. The network evaluation demonstrates that the IoT-based system maintains reliable communication, with minimal latency and packet loss, ensuring smooth face recognition and locker control operations. The throughput results indicate that the network can handle the data transmission requirements of the facial recognition process without significant performance degradation. Furthermore, stable communication performance contributes to faster authentication responses and enhances the overall reliability of the smart locker security system in real-world applications.

Table 4. Network Communication Quality

Parameter	Value
Latency (ms)	120 ± 15
Packet Loss (%)	1.5
Throughput (kbps)	200 ± 10

Table 4 presents the network communication quality of the developed system based on three main parameters: latency, packet loss, and throughput. The average latency recorded was 120 ± 15 ms, indicating that the communication delay between devices and the network server was relatively low and stable, which is suitable for real-time monitoring and authentication systems. The packet loss value of 1.5% shows that only a small portion of transmitted data packets failed to reach their destination, demonstrating reliable network communication performance. In addition, the throughput value of 200 ± 10 kbps indicates that the system could transmit data efficiently at a stable rate during operation. Variations in latency and throughput may have been influenced by network traffic conditions, signal strength, or hardware processing capabilities. Overall, the obtained results confirm that the communication network used in the system provides adequate stability, reliability, and efficiency for practical IoT-based security applications. Furthermore, the low packet loss and consistent throughput values ensured that facial recognition data and authentication commands were transmitted without significant interruption. These findings demonstrate that the proposed IoT architecture can support continuous and dependable smart locker operations, even under varying network conditions.

4. CONCLUSION

This study successfully implemented a YOLO-based facial recognition system for authentication purposes and evaluated its performance under different distances and lighting conditions. The experimental results demonstrated that the YOLO algorithm achieved a a face detection success rate of up to 92% across all test scenarios, confirming its capability to reliably detect facial features in both bright and low-light environments. The detection error remained relatively low, ranging from 2.60% to 4.30%, indicating stable detection performance despite variations in environmental conditions. The authentication tests showed that the system performed most efficiently at 30 cm, where response times ranged from 1.47 s to 2.56 s with error rates below 1.2%. As the distance increased to 100 cm, the authentication delay increased significantly to approximately 9–10 seconds, accompanied by higher error rates. These findings indicate that distance has a greater influence on authentication performance than lighting conditions, as facial features become less distinct when captured from farther distances. Overall, the results confirm that YOLO is an effective and robust algorithm for real-time face detection and authentication applications. Its ability to maintain consistent detection performance under varying lighting conditions makes it suitable for smart security systems such as locker access control. For optimal performance, users should position themselves closer to the camera, while future improvements may focus on enhancing recognition accuracy and reducing response time at longer distances through dataset expansion, image preprocessing techniques, and model optimization.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the institution for providing the facilities, resources, and research environment that enabled the completion of this study. Special thanks are extended to the laboratory staff for their technical assistance during the development, integration, and testing of the IoT-based locker lock security system utilizing YOLO facial recognition and ESP32-CAM technology. The authors also appreciate the valuable contributions, discussions, and constructive feedback provided by colleagues and peers throughout the research process. Their support and encouragement were instrumental in the successful completion of this work.

REFERENCES

[1] A. Rahman *et al.*, “Smart Keyless Locker Design Using Face Recognition Technology Based on the Internet of Things Jakarta Global University Classroom,” *J. Glob. Eng. Res. Sci.*, vol. 4, no. 2, pp. 56–66, Dec. 2025, doi: 10.56904/j-gers.v4i2.162.

[2] A. Choudhary, “Internet of Things: a comprehensive overview, architectures, applications, simulation tools, challenges and future directions,” *Discov. Internet Things*, vol. 4, no. 1, p. 31, Dec. 2024, doi: 10.1007/s43926-024-00084-3.

[3] K. N. Sai, D. T. Sunil, and D. M. Eshwarappa, “A comprehensive review of door lock security systems,” *Int. J. Circuit, Comput. Netw.*, vol. 5, no. 1, pp. 12–17, Jan. 2024, doi: 10.33545/27075923.2024.v5.i1a.61.

[4] S. Mohammed and A. H. Alkeelani, “Locker Security System Using Keypad and RFID,” in *2019 International Conference of Computer Science and Renewable Energies (ICCSRE)*, IEEE, Jul. 2019, pp. 1–5. doi: 10.1109/ICCSRE.2019.8807588.

[5] A. Ramzan, W. Farhan, I. Malahat, and N. Afzal, “Double-Layered Authentication Door-Lock System Utilizing Hybrid RFID-PIN Technology for Enhanced Security,” in *MTME 2025*, Basel Switzerland: MDPI, Aug. 2025, p. 19. doi: 10.3390/materproc2025023019.

[6] R. Keote, S. S. Dandale, P. A. Bhagat, and M. Keote, “Biometric & GSM based Security System for Bank Lockers,” in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, IEEE, Jun. 2024, pp. 1–5. doi: 10.1109/ICCCNT61001.2024.10725225.

[7] N. EL Fadel, “Facial Recognition Algorithms: A Systematic Literature Review,” *J. Imaging*, vol. 11, no. 2, p. 58, Feb. 2025, doi: 10.3390/jimaging11020058.

[8] M. L. Ali and Z. Zhang, “The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection,” *Computers*, vol. 13, no. 12, p. 336, Dec. 2024, doi: 10.3390/computers13120336.

[9] R. Sapkota *et al.*, “YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series,” *Artif. Intell. Rev.*, vol. 58, no. 9, p. 274, Jun. 2025, doi: 10.1007/s10462-025-11253-3.

[10] A. Vijayakumar and S. Vairavasundaram, “YOLO-based Object Detection Models: A Review and its Applications,” *Multimed. Tools Appl.*, vol. 83, no. 35, pp. 83535–83574, Mar. 2024, doi: 10.1007/s11042-024-18872-y.

[11] A. A. Murat and M. S. Kiran, “A comprehensive review on YOLO versions for object detection,” *Eng. Sci. Technol. an Int. J.*, vol. 70, p. 102161, Oct. 2025, doi: 10.1016/j.jestch.2025.102161.

[12] D. Díaz-Delgado, S. M. Vigil-Ramírez, L. J. Acho-Cachay, B. R. Tuanama-Chávez, and L. A. Rojas-Puerta, “IoT-Based Smart Lock with Real-Time Person Detection Using YOLOv5 and Mobile App Integration,” *Rev. Científica Sist. e Informática*, vol. 5, no. 2, p. e1005, Jul. 2025, doi: 10.51252/rcsi.v5i2.1005.

[13] G. C. Preethi, B. S. C. Sekhar, D. Venkatesh, S. Chanikya, K. Krishna, and T. Daniya, “Face recognition for smart door lock system using machine learning algorithms,” 2025, p. 020048. doi: 10.1063/5.0279225.

[14] R. Joshi, R. S. Somesula, and S. Katkoori, “Empowering Resource-Constrained IoT Edge Devices: A Hybrid Approach for Edge Data Analysis,” 2024, pp. 168–181. doi: 10.1007/978-3-031-45878-1_12.

[15] F. Ben Aicha, “Optimized embedded AI: efficient implementation of CNNs on ESP32-CAM for real-time image classification,” *Computing*, vol. 107, no. 11, p. 206, Nov. 2025, doi: 10.1007/s00607-025-01559-z.

[16] A. Maier, A. Sharp, and Y. Vagapov, “Comparative analysis and practical implementation of the ESP32 microcontroller module for the internet of things,” in *2017 Internet Technologies and Applications (ITA)*, IEEE, Sep. 2017, pp. 143–148. doi: 10.1109/ITECHA.2017.8101926.

[17] J. Zhang and N. Hu, “Accuracy and robustness evaluation of deep learning algorithms in facial recognition systems,” *Syst. Soft Comput.*, vol. 7, p. 200252, Dec. 2025, doi: 10.1016/j.sasc.2025.200252.

[18] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, “A Review of Yolo Algorithm Developments,” *Procedia Comput. Sci.*, vol. 199, pp. 1066–1073, 2022, doi: 10.1016/j.procs.2022.01.135.

[19] T. Diwan, G. Anirudh, and J. V. Tembhurne, “Object detection using YOLO: challenges, architectural successors, datasets and applications,” *Multimed. Tools Appl.*, vol. 82, no. 6, pp. 9243–9275, Mar. 2023, doi: 10.1007/s11042-022-13644-y.