Face Recognition Door Lock System with RISC-V

Wagner Pedrosa
CTM Summer Internships 2023
Optical and Electronic Technologies

Paulo Fidalgo
CTM Summer Internships 2023
Optical and Electronic Technologies

***Abstract*—In recent times, the rapid growth of Internet of Things (IoT) applications has sparked a surge of interest in various fields. In line with this trend, this project aimed to create a Face Recognition Door Lock System using an ESP32-CAM, which is a chip with a microcontroller and Wi-Fi, making it great for IoT applications. Adding face recognition enhances the security of the door lock. An interesting aspect of this work is the use of RISC-V, an open-source instruction set architecture known for its simplicity, flexibility, and expandability. RISC-V allows customization and helps achieve better performance while maintain lower power consumption. The research explored the integration of hardware and software to find innovative solutions for smart security applications.**

I. Introduction

In the fast-paced digital era, the adoption of Internet of Things (IoT) applications is on the rise, bringing about positive changes in the lives of individuals globally. Notably, IoT platforms, particularly those geared towards security and surveillance purposes, are projected to expand at an average annual growth rate of 16.9% between 2022 and 2028 [1]. This anticipated growth underscores the increasing role of IoT in shaping various industries and enhancing everyday experiences.

In light of this, and within the context of INESC TEC CTM Summer Internships 2023, an innovative project entitled "Face Recognition Door Lock System with RISC-V" was undertaken with the aim of transforming a traditional door lock into a cutting-edge IoT-based security solution. Leveraging the powerful ESP32-CAM microcontroller and a proficient face recognition algorithm, this project sought to create a versatile, robust, reliable, and energy-efficient doorway security system accessible to all.

The development process encompassed a meticulous analysis of hardware and software integration, providing invaluable insights into achieving optimal efficiency in IoT applications. The main objectives of the project include designing and constructing the circuit of the system, integrating a face recognition algorithm to enable secure access control, developing a user-friendly interface for easy configuration and management, and thoroughly testing and documenting the system to ensure reliability and future enhancements.

This article delves into the project's significant components, starting with the hardware utilized, including the ESP32-CAM microcontroller, FTDI board, jumper wires, relay module, solenoid lock, and transformers. Additionally, it examines the face recognition algorithm deployed, showcasing its functionality and efficiency. The article also explores the development of a mobile app in Swift, serving as the system's user interface for effortless control and management.

Furthermore, the results section delves into the face recognition error rates and energy consumption analysis, providing valuable insights into the system's performance and sustainability. A SWOT analysis evaluates the project's strengths, weaknesses, opportunities, and threats, shedding light on its viability and scalability.

Overall, this article offers a comprehensive understanding of the "Face Recognition Door Lock System with RISC-V" project, demonstrating its potential to revolutionize traditional door security with advanced IoT technology and user-friendly features.

II. Developed work

A. Circuit Implementation

To ensure the successful implementation of the project, a circuit was constructed (Figure 1) where the ESP32-CAM served as the main processing unit. To power the ESP32-CAM, a 5V power supply was employed. This power supply converted the standard 100-240V AC power from an electrical socket into the required 5V DC power that the ESP32-CAM needed to function.

Since the ESP32-CAM lacks a built-in USB port, a separate FTDI board was used. The FTDI board acted as a converter, transforming the USB signal into a format the microcontroller could understand (UART signal). This setup allowed code uploading and provided the necessary 5V DC power for the ESP32-CAM.

The ESP32-CAM was chosen as the central processing unit due to its cost-effectiveness and powerful capabilities. Developed by Espressif Systems, it is based on the ESP32 chip, which incorporates Wi-Fi, Bluetooth (including BLE), and an OV2640 camera module. This combination makes it ideal for various IoT projects. The ESP32-CAM houses two high-performance 32-bit LX6 CPUs and operates on a 7-stage pipeline architecture based on the RISC-V instruction set. It serves as the brain of the project, processing all the information.

Additionally, the ESP32-CAM was used as a controller for the relay, which acted as a mediator between the high-power solenoid circuit responsible for operating the door lock and the low-power ESP32-CAM circuit. By utilizing a low-voltage (5V) and low-current signal from the chip, the relay module could handle the high-voltage (12V) and high-current requirements of the door lock, effectively operating as a virtual

switch, controlling whether the door was locked or unlocked based on commands from the microcontroller.

Furthermore, the solenoid lock was powered by a 12V power supply, which converted the standard 100-240V AC from an electrical socket into the required 12V DC 500mA necessary to operate the circuit. When energy is applied, it generates a magnetic field that pulls the slug in, unlocking the door. Conversely, when the power supply is disconnected, the magnetic field weakens, allowing the slug to move outward and securely locking the door. Importantly, the electronic lock operates efficiently, consuming no power when in locked mode.

In the construction of the circuit, jumper wires were used to connect all the components together, ensuring a functional and well-integrated setup.

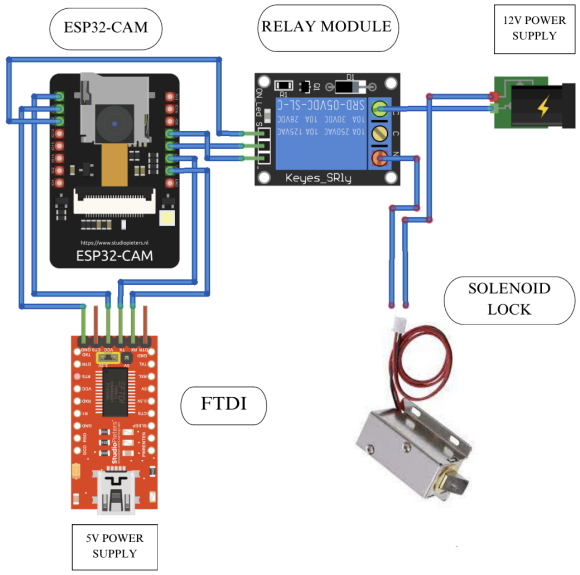


Fig. 1: Circuit Diagram

To ensure proper operation of the Face Recognition Door Lock System, essential connections were established between the ESP32-CAM microcontroller and the FTDI board. These connections serve critical functions to facilitate power supply and bidirectional data communication. Firstly, the 5V pin of the ESP32-CAM was meticulously connected to the VCC pin of the FTDI board. This linkage ensures the provision of the necessary 5V DC power from the FTDI board to the ESP32-CAM, enabling its smooth functioning and operation. Secondly, the establishment of a common ground reference is vital for operation. Consequently, the ESP32-CAM's GND pin was appropriately connected to the FTDI board's GND pin.

Furthermore, to enable data exchange with external devices, the ESP32-CAM's Universal asynchronous receiver/transmitter (UART) pin (named U0R) was connected to the FTDI board's TX (Transmit) pin. This connection allows data transmitted by the ESP32-CAM to be accurately received by the FTDI board, facilitating effective communication with external peripherals.

Additionally, to enable the ESP32-CAM to receive data from the FTDI board, the ESP32-CAM's UART (U0T) pin was connected to the FTDI board's RX (Receive) pin. This bidirectional communication enables data processing and coherent interaction between the ESP32-CAM and external devices.

The integration of a 12V solenoid lock circuit required the use of a relay. To ensure the system's full functionality and effectiveness, critical connections between the ESP32-CAM and the relay were established.

Firstly, a connection was made between the 5V pin of the ESP32-CAM and the VCC pin of the relay. This ensures the relay receives the necessary 5V DC power for its seamless operation as an integral part of the circuit. The relay's reliable power supply is essential for its efficient functioning in securely managing the door lock mechanism.

Secondly, to ensure a robust signal communication link between the ESP32-CAM and the relay, a common ground reference was carefully established. This connection involves linking the GND pin of the ESP32-CAM with the GND pin of the relay.

For efficient control of the relay and precise door lock management, an essential connection between the ESP32-CAM's IO12 pin and the relay's IN0 pin was made. This connection enables the ESP32-CAM to send control signals to the relay. By activating the IO12 pin with specific high or low signals, the ESP32-CAM instructs the relay to switch its state, effectively controlling the door lock's locking and unlocking actions.

In conclusion, the described connections form a well-designed and integrated system for the Face Recognition Door Lock. By establishing these connections correctly, the system can function smoothly, ensuring reliable power supply, effective signal communication, and precise and secure control of the door lock mechanism.

The details regarding the connections between the ESP32-CAM and the FTDI, as well as the connections between the FTDI and the relay, are presented in Table I and Table II, respectively.

ESP32-CAM	FTDI
5V	VCC
GND	GND
UOR	TX
UOT	RX

TABLE I: ESP to FTDI Connection Mapping

ESP32-CAM	Relay
5V	VCC
GND	GND
IO12	IO0

TABLE II: ESP to Relay Connection Mapping

B. Algorithm Pipeline

To accomplish the complex task of correctly recognizing the face of a person authorized, a library built by Expressif Systems, the creator of the ESP32-CAM chip, was used. Entitled *FRMN*, this lightweight Human Face Recognition Model, was built around the architecture of *MTMN* (*MobileNetV2* [2] and *Multi-task Cascade Convolutional Network* [3]), using the *ArcFace* loss function [4]. Although those algorithms are not state of the art in the field of face recognition, they

provide a balanced trade-off between accuracy and efficiency, particularly suitable for embedded systems which are built to have low computational power, leading to a reduced power consumption.

The face recognition process begins with capturing a 320x240 pixels resolution image, in RGB565 format, using the embedded camera of the ESP32-CAM. The captured image is then converted to RGB888 format and subsequently, subjected to three stages of convolutional networks, enabling the detection of facial coordinates and landmarks, such as eyes, nose, and mouth [3]. The algorithm is configured to detect a single face, identifying the box with the highest probability of being a face [3]. Once the face is detected, it undergoes alignment using the landmark coordinates, resulting in a new 56x56 pixels resolution image, tailored for the face recognition algorithm. A distinctive face identification (ID) is constructed based on the aligned image, allowing for individual recognition. To compare the new face ID with existing ones, the cosine distance between them is calculated [4]. A critical aspect of the process lies in the threshold, set at 0.55. If the similarity between the newly generated face ID and any of the previously stored IDs exceeds this threshold, the person is accepted as a match, and the door opens for five seconds, otherwise it stays closed. This refined approach to face recognition ensures an efficient and accurate identification process, enabling secure and reliable access control for embedded systems powered by the ESP32-CAM.

The process is similar when the user intends to add a new Face ID to the system. In this scenario, instead of using only one image, ten samples are taken to ensure a more reliable and accurate representation of the person's face. By incorporating multiple samples, any potential variations in facial expressions and angles are considered, enhancing the system's ability to capture the individual's facial features comprehensively. After successfully capturing ten samples, a new face ID is generated and then stored in memory as an authorized person to open the door.

The pipeline of the process is shown in Figure 2.

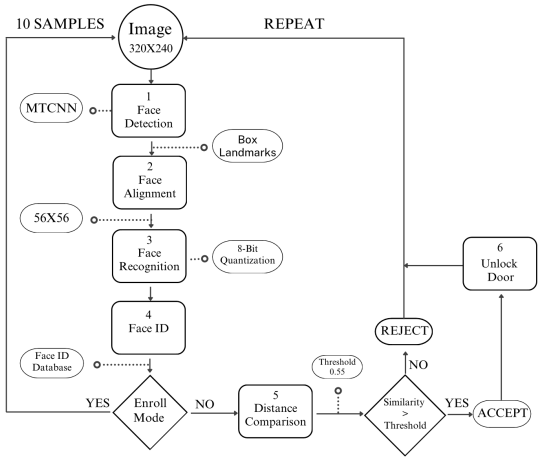


Fig. 2: Algorithm Pipeline

C. Mobile App

Using Swift, an iOS application, was developed with the primary aim of enabling efficient management and control of the entire system.

The app communicates with the microcontroller via Wi-Fi, utilizing the TCP/IP protocol stack for managing communications. In this case, the microcontroller is used as an Access Point, requiring users to connect their mobile phones to the microcontroller and enter the password to gain control. The app operates through the transmission of messages to the ESP32-CAM, which, in turn, interprets these messages to execute the appropriate actions.

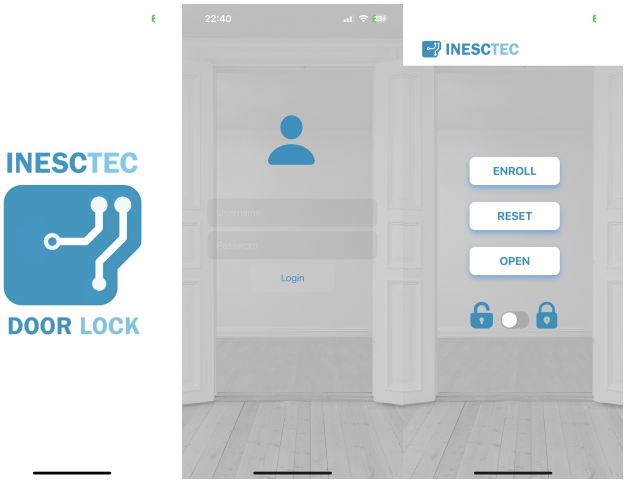
In order to enhance and optimize the user experience (UX), the application employs a visually appealing **splash screen**, as demonstrated in Figure 3a, characterized by a seamless combination of blue and white tones converging in order to create the INESC TEC logo. The strategic use of the blue color instills a perception of trust and security, while the white color contributes to a clean and uncluttered design. It is pertinent to mention that the application's intrinsic simplicity ensures a high level of user-friendliness, catering to individuals of diverse age groups and technology literacy.

To fortify security even further, a **login page**, as shown in Figure 3b, follows the splash screen, obliging users to provide their unique Username and Password. Access is granted solely upon successful validation of both credentials against the system's stored records.

Upon successful login, users are directed to the **main screen**, a clean page, as presented in Figure 3c, with four named buttons and a well-placed background, ensuring easy navigation and reducing confusion. Each button is named according to its main functionality, keeping their labels short and clear. Among these buttons, the "**Enroll**" button enables users to effortlessly register a new ID by simply pressing it. Upon activation, the camera takes ten photos, and once completed, the person gets enrolled in the system, and their information is saved in memory, eliminating the need of repetitive registration. Conversely, the "**Reset**" button gives users an efficient way of deleting all existing IDs stored in memory. Furthermore, the "**Open**" button grants access without the necessity of verifying if the person in front of the camera is a valid ID. This feature is ideal for welcoming visitors without recording them as genuine IDs.

Finally, the "**Lock/Unlock**" toggle button. When in "**Lock**" position, the system becomes impenetrable, providing the user with exclusive control to override the face ID algorithm and restrict access to everyone, regardless of whether the "Open" button is pressed. This restriction ceases when the "**Unlock**" button is activated, restoring the regular functions. This safeguard is particularly useful when the owner wants to ensure the door is inaccessible, overnight or when no one is at home. To sum up, this iOS app has been designed with the primary goal of controlling the circuit, combining advanced technology with a user-friendly interface. The system ensures a smooth and reliable experience, providing users with the

convenience of managing the ESP32-CAM effectively.



(a) Splash Screen (b) Login Page (c) Main Screen

Fig. 3: Mobile App

III. Results

A. Face Recognition Algorithm

The project focused on developing a face recognition-based door lock system, with particular emphasis on evaluating the model's performance and error rate. For a first approach, two potential errors were identified: granting access to impostor and denying access to enrolled individuals. To assess these aspects, a comprehensive test was conducted involving 20 distinct identities, half of whom were enrolled users, and the other half acted as intruders. To ensure the validity of the results, and, at the same time, simulate a real application of the door lock system, the tests were conducted in a controlled environment, with consistent lighting conditions, background settings and distance from the camera sensor. This approach minimized external factors that could influence the algorithm's performance, enabling a reliable assessment of its accuracy and efficiency. By simulating real-world scenarios, it was possible to obtain accurate insights into the system's ability to distinguish between genuine users and unauthorized intruders, enhancing the system's overall effectiveness and reliability for practical use.

During the tests, each individual was randomly presented three times in front of the came, its face was detected, and the algorithm generated similarity values (from 0.00 to 1.00) for comparison. These similarity scores were aggregated in intervals of 0.05 and plotted on a graph versus frequencies. The idea of this test was to have two distinct, separated and well-defined distributions, one for enrolled IDs and the other for intruders. Ideally, enrolled individuals similarity values would be clustered near 1.00, while impostor scores should predominantly be close to 0.00. Additionally, the totality of enrolled individuals similarity values should be to the right of the chosen threshold, set at 0.55, signifying accurate identification. Impostor scores would preferably cluster to the left of the threshold, indicating an inability to unlock the door. The error of the algorithm can be measured by the overlap between authorized and unauthorized distributions, indicating potential issues in distinguishing genuine users from impostor, which poses a threat to the system's security.

The results presented in Figure 4 revealed that both genuine and impostor distribution were close, indicating a potential threat and weakness in the algorithm. Analyzing the impostor side, the maximum and minimum similarity values were 0.45 and 0.20, respectively, with a predominant emphasis on the 0.35 - 0.40 similarity interval, where 15 samples fell. This indicates that impostor scores were relatively near the threshold. On the positive side, there were no cases of false positives, meaning there were no instances of zero effort forgery, which is crucial for a security system.

However, while there was a well-defined distribution in the impostor side, the genuine distribution did not present any predominant interval, highlighting a weakness in the algorithm's discriminatory capabilities. The maximum and minimum genuine similarity values were 0.88 and 0.45, respectively, with 2 values falling below the defined threshold, representing the false negative error.

Overall, despite conducting 60 tests, there were only two false negatives and no false positives in the evaluation, indicating promising performance in terms of minimizing false positives. However, the algorithm's weakness in accurately distinguishing genuine individuals suggests the need for further improvement to minimize false negatives and enhance the system's overall performance. In addition, more test should be conducted before implementation of the system in a real world scenario.

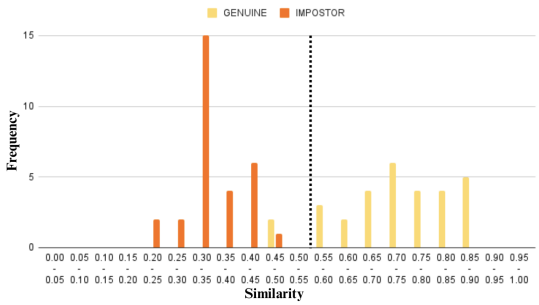


Fig. 4: Impostor and Genuine similarity distributions

B. Energy Consumption

The application of IoT to control the door offers numerous advantages, but it comes with the need for constant power supply with the aggravating factor of using Wi-Fi to communicate. In this context, measuring energy consumption becomes crucial, particularly the impact of Wi-Fi usage on overall power requirements. This analysis aimed to evaluate the power consumption of an ESP32-CAM based door control system through three different tasks, considering the energy used with and without Wi-Fi connectivity. The experiment involved measuring current and voltage in the ESP32-CAM

circuit using a multimeter in series, and calculating power using the power equation:

$$P = V \times I$$

where **P** stands for power (in Watt), **V** for Voltage (in Volt) and **I** for Current (in Ampere). Three distinct tasks were evaluated: idle system operation, face detection and image processing (taking around 8 seconds from the moment a face is presented in front of the camera), and door opening, which required powering the door lock circuit for 5 seconds (if the person is accepted). It is worth noting that the voltage value of the ESP32-CAM circuit maintained constant, 5.17V.

The results are shown in Figure 5. The most significant finding was that Wi-Fi usage results in approximately 1W of additional power consumption across the different tasks. Furthermore, the energy consumption ratio with Wi-Fi and without Wi-Fi decreased as tasks became more complex, demonstrating that the cost of using communications dissipated as the complexity of the work increased. The idle system operation (only powered) exhibited a 21 times higher energy consumption when using Wi-Fi compared to without it. During image processing, this ratio reduced to 4 times, and during door opening, it further decreased to 2 times, considering the 5-second duration the door remained open. Additionally, when the door opens, the lock circuit consumes 6W extra for 5 seconds, to operate the solenoid lock (12V 500mA).

The analysis of energy consumption in the ESP32-CAM based door control system highlights the notable impact of Wi-Fi usage on overall power requirements. This insight is valuable for optimizing power management strategies, particularly in minimizing Wi-Fi usage during idle periods and prioritizing power-intensive tasks when needed. By understanding the energy consumption patterns, it is possible to enhance the system's efficiency and sustainability, ultimately contributing to the advancement of IoT applications in smart technologies.

	No task	Processing image	Opening the door
Without WI-FI	0.052W	0.312W	0.724W
With WI-FI	1.086W	1.293W	1.450W
Ratio	21	4	2

Fig. 5: Impact of Wi-Fi on power in different tasks

C. Demonstration Video

A video was created to show all the implemented capabilities of this project. It is possible to watch via the following link: https://youtu.be/BVR8iSGyL5U

D. SWOT Analyses

A SWOT analysis was crucial to evaluate the project's strengths, weaknesses, opportunities, and threats.

Firstly, the project's key **strength** lies in its utilization of the ESP32-CAM, which features a well-regarded RISC-V processor - an open-source instruction set architecture (ISA). RISC-V offers various advantages, including accessibility, collaboration, customization, and no licensing costs. Its transparent design fosters trust and encourages a diverse community of developers to collaborate globally. Furthermore, this open nature promotes innovation, research, and education while facilitating the growth of a robust ecosystem of hardware and software tools. Moreover, the project's cost-effectiveness, ranging from 30 to 40€, showcases its affordability. The setup process is straightforward due to the availability of all necessary components and easy face recognition code implementation. Another advantage is the user-friendly mobile app that enables effortless operation with clear and intuitive buttons. Despite utilizing Wi-Fi, the system demonstrates commendable energy efficiency, with an average power consumption of 1.086W (considering that most of the time the system is in idle), minimally impacting overall power usage compared to other household appliances. Lastly, the project exhibits strong expandability, designed to accommodate future enhancements and additional features seamlessly, enhancing its potential for further development and scalability.

However, being a system on a chip, the project faces challenges, representing its **weaknesses**, such as the limited computational power and camera quality, impacting the implementation of a sophisticated face recognition algorithm, leading to compromised performance and accuracy. Additionally, the 4GB memory capacity poses constraints on storing numerous IDs, limiting the system's capabilities to be installed in a place where a large number of identities would have access.

To enhance the system, several **opportunities** can be pursued. Primarily, improving the face recognition algorithm is crucial to achieve better accuracy and establishing a secure and reliable system, making it suitable for real-world applications. Furthermore, integrating a smart home application and incorporating a live-stream platform and notification page into the mobile app will provide users with precise control, enabling access to live camera feeds and real-time updates. Exploring the adoption of solar cells for powering a battery, that would act as a power backup for the entire system and implementing an economic mode to optimize energy efficiency will promote sustainability and reduce dependence on traditional power sources. By seizing these opportunities, the face recognition system can evolve into a more advanced and practical solution, expanding its potential applications in various industries, including security and smart surveillance.

Lastly, identified **threats** revolve around security risks, emphasizing the crucial need for a secure and robust algorithm, to prevent forgery and withstand attacks. Moreover, the system's vulnerability to energy failures poses a significant concern, potentially making the door inaccessible even to enrolled users. Lastly, the dependence on Wi-Fi connectivity for the mobile app creates usability issues when network access is unavailable, affecting the app's performance while

the face recognition control remains operational. Addressing these concerns is vital to ensure the system's effectiveness and user satisfaction.

IV. Conclusion

In conclusion, this project successfully demonstrated the development of a versatile and efficient IoT-based security solution for door access control, maintaining simplicity, flexibility, and low power consumption. While the system exhibited commendable performance, continuous research and development are essential to enhance accuracy, security, and usability, paving the way for wider adoption in smart security applications.

References

[1] "Smart Home Platforms Market Size Report, 2022 - 2028". https://www.grandviewresearch.com/industry-analysis/smart-home-platforms-market-report (accessed on Aug. 04, 2023).

[2] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. MobileNetV2: Inverted Residuals and Linear Bottlenecks, March 2019. arXiv:1801.04381 [cs].

[3] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. *IEEE Signal Processing Letters*, 23(10):1499–1503, October 2016.

[4] Jiankang Deng, Jia Guo, Jing Yang, Niannan Xue, Irene Kotsia, and Stefanos Zafeiriou. ArcFace: Additive Angular Margin Loss for Deep Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):5962–5979, October 2022. arXiv:1801.07698 [cs].