Computer Science and Information Systems 21(4):1359–1387 https://doi.org/10.2298/CSIS240229038B

Biometric Lock with Facial Recognition Implemented with Deep Learning Techniques

José Misael Burruel-Zazueta[1], Héctor Rodríguez-Rangel[1], Gloria Ekaterine Peralta-Peñuñuri[1], Vicenç Puig Cayuela[2], Ignacio Algredo-Badillo[3], and Luis Alberto Morales-Rosales[4,⋆]

[1] Tecnológico Nacional de México/Instituto Tecnológico de Culiacán
Juan de Dios Batiz No. 310 pte, Guadalupe, 80220 Culiacán Rosales, Sin., México
d13170395@culiacan.tecnm.mx
hector.rr@culiacan.tecnm.mx
gloria.pp@culiacan.tecnm.mx
[2] Institut de Robòtica i Informàtica Industrial, CSIC-UPC
C/ Llorens i Artigas 4-6, 08028 Barcelona, España
vicenc.puig@upc.edu
[3] CONACYT-Instituto Nacional de Astrofísica, Óptica y Electrónica
Tonantzintla 72840, Puebla, México
algredobadillo@inaoep.mx
[4] CONACYT-Universidad Michoacana de San Nicolás de Hidalgo
Avenida Francisco J. Múgica S/N Ciudad Universitaria, Morelia, México
lamorales@conacyt.mx

Abstract. The increased criminal activity associated with unauthorized entry into facilities has become a global concern. Traditional mechanical locks suffer from drawbacks such as key loss, theft, duplication risks, and time-consuming operation. Therefore, biometrics has been explored as a key to accessing a restricted area. However, some challenges still need to be solved in developing such systems, including user registration, response speed, maintainability, and the ability to distinguish between real and fake individuals. This paper proposes and develops a biometric lock system (BLS) whose opening is performed by recognizing a person's face. It solves the challenges of re-training, antispoofing, real-time response, and works inside an embedding system. The proposed BLS overcomes these challenges using a pre-trained network called FaceNet for feature extraction and coding into 128-dimensional vectors. We use the characteristic vector provided by FaceNet and a cosine distance to recognize the persons. It also incorporates ResNet18 + remote photoplethysmography (rPPG) to avoid spoofing. The architecture was implemented in a BLS, demonstrating an impressive false acceptance rate of 0% under varying lighting conditions, with an average response time of 1.68 seconds from facial detection to door opening. The BLS has easy maintainable devices, providing enhanced security by accurately identifying individuals and preventing unauthorized access. The system can distinguish between real and fake people without using specialized hardware. Making it a versatile solution suitable for homes, offices, and commercial spaces. The results underscore the potential efficacy of our proposed BLS in mitigating security challenges related to unwarranted access to restricted facilities.

Keywords: FaceNet, Jetson Nano, CNN, Door lock, Embedded system

⋆ Corresponding Author

1. Introduction

Building burglary is a security issue that remains one of the main problems in the Americas according to worldwide incidence rates [40]. Only some countries, such as the United States, Mexico, and Canada, have reduced their rates in the last 15 years. However, cases such as Brazil, Ecuador, Guatemala, and Chile have alarmingly increased crime rates. Some locks are used regularly to protect buildings, houses, and offices, which are activated mechanically, electronically, smartly, and hybrid.

The facility or asset protection is done according to its importance (i.e., the bank deposit lock differs from a building lock). The type of lock is defined by the asset's characteristics, such as the security level, user number, and response time. The scenarios may vary depending on these concepts since the security levels for protecting a bank vault are higher than for a house or office. Using the same example, the number of users allowed in both restricted areas varies.

Mechanical locks boast resilience in power outages and durability owing to their robust materials. Nevertheless, they rely on physical keys susceptible to loss, duplication, and time-consuming operation, presenting challenges during emergencies or high-traffic scenarios. Conversely, electronic locks, including RFID cards, passwords, cryptography, NFC, biometrics, and hybrid systems, offer keyless entry, remote monitoring, and swift response times but are rendered useless without electricity and hinge asset security on the safety of accompanying devices or password management, often incurring higher costs. RFID cards [46] emulate traditional keys with enhanced user-friendliness and speed but remain vulnerable to loss or theft. Common passwords bolster security but heighten the risk of forgetting access credentials. Advanced techniques like one-time passwords (OTP) [43], IoT [43], cryptographic solutions [8], and NFC devices [42] necessitate additional devices for code or signal generation, elevating security levels but increasing user complexity and usability challenges in having to manipulate these devices.

Cost of the approaches related to smart locks are based on biometrics. These locks use the features present in the user's body as an access key [28]. Several user's body features distinguish one from the other, e.g., fingerprints [33, 45, 66], eye iris [44, 57, 58], vein patterns [54,62], voice [13,27,65], face [6,7,17,47], among others. These locks are easy to use, have a high response speed, and do not require additional devices or keys.

A hybrid lock consists of two or more unlocking techniques to increase the security of the systems [18, 23, 31]. They are used mostly in restricted areas, such as safes, bank vaults, and militarized areas. They are characterized by being difficult to use and prolonged access response.

There are some differences with respect to the biometric features implemented in a lock. Facial recognition requires less user cooperation with the system. Unlike fingerprint, vein, or voice recognition, the user has minimal interaction with the system [5]. The fingerprint or vein recognition requires the user to directly contact a scanner for the biometric readings. Alternatively, for an iris reading to be effective, the user must position his/her eye in proximity to an iris scanner. Therefore, showing a face to a camera is a simple task that almost anyone, including children, can do.

Face recognition techniques can be divided into three approaches: local, holistic, and hybrid [32]. Local approaches [2, 50] only use some features present on the face (eyes, mouth, nose, etc.), generally having faster processing speed but poor recognition effectiveness. Holistic approaches [9, 20] use the entire face for feature extraction, with high effectiveness and speed with adequate hardware. Hybrid approaches [11, 37] combine the benefits of holistic and local approaches. However, they are slow to process but highly effective.

The most widely used technique in the field of face recognition is convolutional neural networks (CNNs) [1]. Due to their high specialization in image processing, they are a holistic approach that can represent facial features in small information vectors. The main disadvantage of CNNs is the high demand for data since they require high training to reach high rates of recognition effectiveness.

This training can take many hours of work depending on the number of users to be registered. This process presents a disadvantage since the administrator must stop the lock operation with each new user registration to train the facial identification system until the process finishes. In addition, the registration images of users to a facial recognition system are limited by the number of possible samples to be captured during the registration. This means that taking a few thousand pictures to train a CNN model is not feasible since this process would be unaffordable for the administrator and the users to register. This constraint is related to *One-shot learning*, which refers to a model that needs to be learned to classify with only one opportunity [38] (or shot).

Facial recognition locks can also be spoofed by users who present people's printed or digital photographs in front of a camera, making another challenge in developing facial recognition biometric locks. However, there are few proposals to deal with this type of attack. Some of these techniques directly involve the user by requesting some movement to verify a real person (e.g., blinking [14]).

Others use thermal cameras [64] with simple convolutional neural networks trained to identify real faces [10]. These techniques have shortcomings in that they can increase hardware costs (thermal cameras), can be spoofed by videos (blinking), or the light conditions and quality of a photograph could allow intruders to gain access.

In summary, biometric locks based on facial recognition are a viable alternative for building access control due to their effectiveness, response time, and ease of use. For its development, it is necessary to overcome some challenges: the amount of data available for user registration (one-shot-learning), detection that users are real (anti-spoofing) without impacting the ease of use and avoid the use of specialized hardware not to raise costs, and finally implement a recognition system in a maintainable embedded system.

Our proposal takes a strategic two-phase approach to development. The first phase involves coding and integrating face recognition techniques, specifically addressing the challenges of one-shot learning and anti-spoofing. In the second phase, we design a modular system of independent devices that seamlessly work together as an electronic door lock, ensuring a comprehensive and robust solution.

In this work, as the solution for *One Shot Learning*, we selected a CNN architecture named FaceNet [52], which extracts the characteristics of the faces and stores them in a vector, thus compares them to identify people based on their similarities. The main advantage of FaceNet is that it is a pre-trained network, so it does not require numerous images for facial identification. Hence, it can extract the features of new faces with a single image and store them in a vector. Another important feature of FaceNet is the effectiveness demonstrated during its development. According to the results obtained by the authors, it reached 99.67% effectiveness with training of more than 500 million images from more than 10,000 different subjects of all races, ethnicities, and ages.

To avoid spoofing through photographs displayed on the camera, a pre-trained CNN for facial skin detection called ResNet18 [21] and the remote photoplethysmography technique (rPPG) [34] were used. ResNet18 extracts small sections of the image corresponding to the user's facial skin. These skin extracts are processed with an rPPG algorithm, which can identify the heart rhythm through the light reflection changes caused by the blood circulation under the skin. Photographs do not show a change in reflected light, while videos depend directly on the device's luminance, which is why the rPPG functions as an anti-spoofing filter. The ResNet18 is a low-density network that decreases the performance impact on the computer and, in conjunction with a lightweight mathematical algorithm such as rPPG, provides a low computational cost reality detector.

We integrated an anti-spoofing system (rPPG), a feature extractor (ResNet18), and a facial recognizer (FaceNet) in an embedded system to develop a biometric lock with a zero percent false positive rate with a real-time response speed. It is important to mention that the ResNet18 + rPPG + FaceNet offer a balance between effectiveness, speed being executed, and reality detection compared to other prototypes highlighted in the state of the art, which, in some cases, demonstrate high performance in one aspect but too low or no performance in another.

Finally, a biometric-lock system (BLS) capable of controlling access through facial recognition was developed. The BLS was divided into three modules: (i) user interfaces and control, which describe the interaction of the hardware elements (Camera, Jetson Nano, display, and Arduino UNO) that are responsible for running the facial recognition system (FRS) and door opening, as well as interacting directly with the user, (ii) FRS, which represents the software responsible for processing the images received for registration or authentication of users of facial recognition by implementing ResNet18 + rPPG to validate that the faces received belong to real people and FaceNet for the face characteristics extraction, and (iii) Power circuit, this module is responsible for managing the electrical energy to lock and unlock the door.

This three-phase modulation allows us to visualize the elements of the prototype individually. Thus, it is possible to perform individual maintenance on each hardware device separately, increasing the maintainability of the BLS. Although there are BLS alternatives on the market, most of these devices are completely hermetic, making the maintenance task difficult, in addition to their high costs.

The scenario proposed in this work is to generate a lock capable of allowing access to a large number of people, which can vary from the size of a family (3-8 members) to a corporation of workers (100-200 members). It is also necessary not to require any key, password, card, or additional device, as well as to reduce the user's cooperation to only present his/her face to the camera. Therefore, it is considered a real-time response speed (0-2 seconds) to avoid a backlog of people trying to gain access to the building. This time was determined by observing a school RFID access control system under actual operating conditions. Also, a false acceptance rate (also known as false match rate, FMR) under 0.1% according to the regulations of NIST [19].

Therefore, the main contributions of this work are enumerated as follows:

- We propose a novel three-module architecture for facial recognition (ResNet18 + rPPG + FaceNet) available in [5]. ResNet18 and rPPG are responsible for calculating

[5] https://github.com/MisaelZazueta/Face-Recognition-Door-Lock

the visible heart rate on the user's face to rule out those who do not have a heartbeat, thus limiting facial recognition to real faces and not intruders seeking to circumvent the system through photographs or videos. All real faces are passed to FaceNet for identification with a single sample of the person's face; we do not carry out a learning stage; instead, we use the characteristic vector provided by FaceNet and cosine similarity to distinguish and recognize the persons listed in our database. This architecture avoids specialized equipment or further interaction with the user. It also provides a lightweight architecture suitable for running on embedded devices such as the Jetson Nano.

- We designed an interface and control interface architecture that consists of four elements (Jetson Nano, Camera, LCD Display, and Arduino UNO) for the software execution. This architecture is based on deep learning techniques (Jetson Nano) capable of making decisions (Boolean) and sending signals (Arduino UNO) to a power circuit to control the energy flow for an actuator. These interfaces are also responsible for generating the necessary images (Camera) to be processed by the software and interacting with the user as visual support (LCD screen).
- We built a BLS based on facial recognition by integrating the facial recognition system, the user and control interface architecture, and a power circuit. This BLS is suitable for the use suggested in the above-proposed scenario and is maintainable due to its software and hardware modules division.

According to the obtained results, the BLS has a high degree of security. This is because the system did not allow access to people not registered in the system, i.e., it obtained a 0% false acceptance rate (0 false positives). The system had only non-identification errors for some users. That is, it obtained an average false rejection rate 9%. However, the rate of people who achieved at least one successful access was 95.12%. In addition, the successful results showed an effectiveness of 91% in all shots. Finally, the response time from face detection to door opening averaged 1.68 seconds. These results, in conjunction with a 100% effectiveness in antispoofing, demonstrate a balance between the parameters sought for BLS development.

The following sections of this paper are organized as follows. Section 2, Related Work, briefly describes some of the work done by other authors. Section 3 describes the Materials and Methods used to develop the work. Section 4 discusses the results obtained in the BLS testing stages. Finally, Section 5 contains the conclusions inferred from the results obtained.

2. Related Work

The development of biometric door locking systems is a topic that computer researchers are addressing to develop security devices with better quality standards. However, some approaches specialize in hardware or software, and the minority of these developments encompass hardware and software. The following is a summary of some of the work related to solutions proposed by other authors. First, we describe work on biometric locks based on facial recognition, each using different hardware devices and facial recognition techniques to perform their tasks. Afterward, we describe outcomes with facial recognition techniques to visualize an overview of what is currently being used by the scientific community in terms of facial recognition.

In [41], the authors use a Raspberry Pi to run the facial recognition software. Because this device has limitations in running deep learning-based applications, a Neural Compute Stick 2 (NCS2) is added as an auxiliary device. Deep learning-based classification tasks are processed on the NCS2. Also, they use an electromagnetic lock driven by a coil, which in turn responds to the signals emitted by a GPIO output of the Raspberry Pi. A third party developed the implemented facial recognition. It consists of a neural network optimized for microcontrollers called MobileNet [24], combined with a Single Shot Detector (SSD) [36] system to detect that it is a real person. The results correspond to an accuracy of 88.75% in facial recognition with a population of 16 people and five repetitions per person. In addition, the system did not fail in any of the five attempts to deceive the design through digital photographs displayed on smartphones.

The authors of [35] developed a face recognition system implemented on a Jetson Nano microcomputer. A Haar algorithm performs face detection and FaceNet performs face recognition [52]. This system does not mention any reality detection stage to verify that the recognition applies only to real people. A reed switch initiated the algorithm once a person opened the door. Their results showed an accuracy of 97.7%

The paper [26] describes a door lock based on a Raspberry Pi microcontroller. Other hardware elements in this paper are a solenoid in charge of activating and deactivating the door lock, a relay to control the voltage flow to the solenoid, a camera module, and a buzzer that works as an audible alarm against intruders. Facial recognition is based on a simple convolutional neural network (CNN) [49]. The database comprised 1100 images, of which 1040 were for training and 60 for testing. The testing process was performed on five registered people and five not registered. Giving 120 samples at 0.5 m, 1.0 m, and 1.5 m distance and four-light conditions: morning, afternoon, evening, and night. The accuracy in 1.0 m and 0.5 m was 100% in all light conditions. The accuracy obtained at a 1.5 m distance was 97.5% in the morning, afternoon, and evening. At night it was 100%. The average accuracy was 97.83%.

A hierarchical system is developed in [60]. In this work, two neural networks operate together to determine the identity of individuals. A ResNet101 [29, 30] was modified to define the faces from images. Subsequently, the FaceNet [52] extracts a feature vector from the face, which is compared inside the database to identify the person. All this is executed inside a Raspberry Pi 4 microcontroller. The door lock is composed only of a servo-motor that transforms the electrical signals into mechanical movements(i.e., it rotates to drive a sliding door once it receives electrical signals directly from the microcontroller). In addition, this system sends e-mail notifications to the administrator every time someone interacts with the facial recognizer. The accuracy obtained was 87.36%. The tests consisted of a total of 230 samples.

In [15], a facial recognition system is described whose main feature is the inclusion of a stage of reality detection of the person by blinking eyes. This has the limitation that intruders can show videos of authorized persons blinking to the camera. The detection of faces in the images is performed by a technique called Histogram of Oriented Gradients (HOG) [16], which generates a vector of features. This vector is classified by a simple Support Vector Machine (SVM) [22]. The software is run on a Raspberry Pi 3 B+ microcontroller. The power module for the proposed door lock consists of a relay, a solenoid, and a power supply. Furthermore, an ultrasonic sensor is included to detect if a person is in front of the Pi-Camera used to obtain the images. The results generated in this work were

92.68% of maximum accuracy in optimal lighting conditions, and the average execution time was 10.196 seconds.

In [51], two different methods are used for face recognition. The first one is based on a local binary pattern histogram (LBPH) algorithm [55]. Each histogram created by this algorithm represents each image in the database. Then, histograms generated by a real-time video are compared with those in the database. The Euclidean distance between histograms is obtained. The second method uses a Histograms of Oriented Gradients (HOG) algorithm [16], which extracts a 128-dimensional feature vector from each face. Like the previous algorithm, the Euclidean distance between the vectors obtained in real-time and those found in the database is calculated. Both algorithms are run on a Raspberry Pi, which controls a MOSFET switch that activates and deactivates the solenoid that locks the gate. A magnetic sensor tells the system if the door has been closed to relock the solenoid. Also, a system to detect the presence of users nearby is used. This uses an Arduino, a motion sensor (PIR), and a lamp. When the motion sensor sends a signal to the Arduino, it turns on the room lamp. The accuracy of the LBPH method was 98%, while HOG was 96%, in a total of 200 tests for each.

Related work describing the development of a biometric lock based on facial recognition focuses primarily on the interaction of hardware elements and facial recognition software. In the works described above, there is an imbalance in the anti-spoofing issue, since, there are few authors who address this issue to prevent malicious attacks by intruders. In this work, a biometric-lock system is developed that approaches all the necessary aspects for implementing a door security system based on facial recognition, with the advantages of embedded devices, but with an individual focus on facial recognition and an anti-spoofing technique using a Resnet18 + rPPG technique that calculates the heart rate from changes in the user's skin tone.

3. Materials and Methods

Hardware and software complement each other to develop a biometric-lock system (BLS) that allows users to access a restricted area through facial recognition. The development methodology to construct the architecture proposed for the BLS is divided into three modules, as shown in Figure 1. The first module refers to the development of the Facial Recognizer System (*FRS*). It is composed of two blocks: (i) *User Registration* and (ii) *User Authentication*, up to the interaction of the information exchanges that exist between the modules. The second module, named *User and Control Interfaces*, consists of four hardware elements and executes the logic tasks. The third is the electronic module, named *Power Circuit*, for locking and unlocking the access door when receiving a signal. We explain each module in the following subsections.

3.1. Facial Recognizer System FRS

FRS is divided into two main blocks. The *User Registration* block has the purpose of registering the people who can unlock the door to gain access, this task is done by an administrator. On the other hand, the *User Authentication* block is in charge of facial

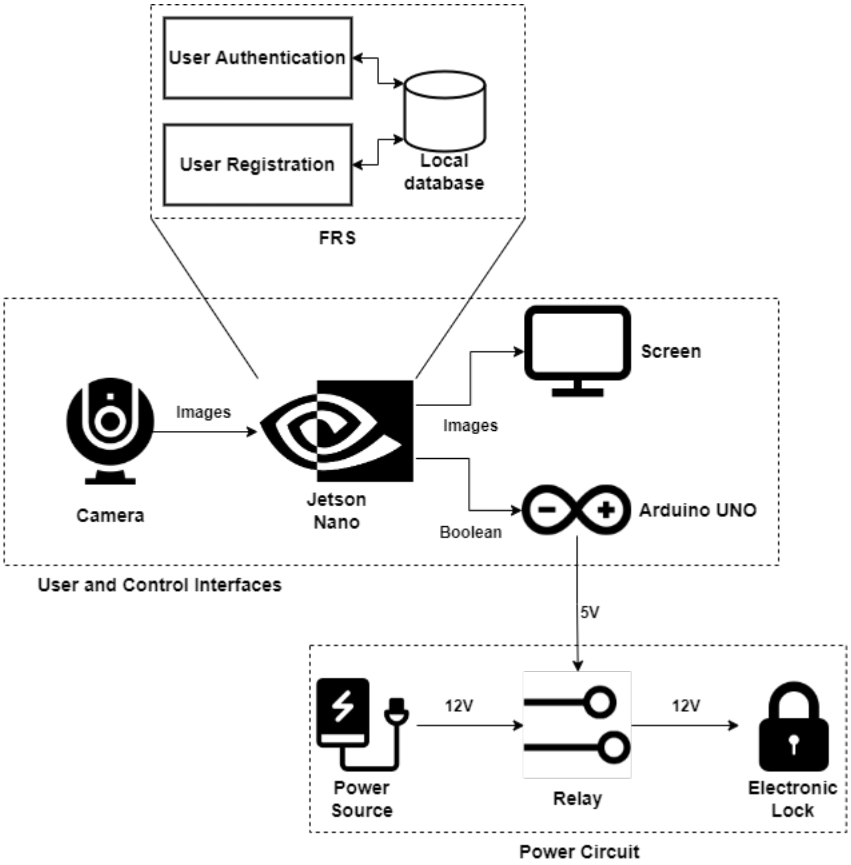


Figure 1. Development methodology to construct the BLS architecture.

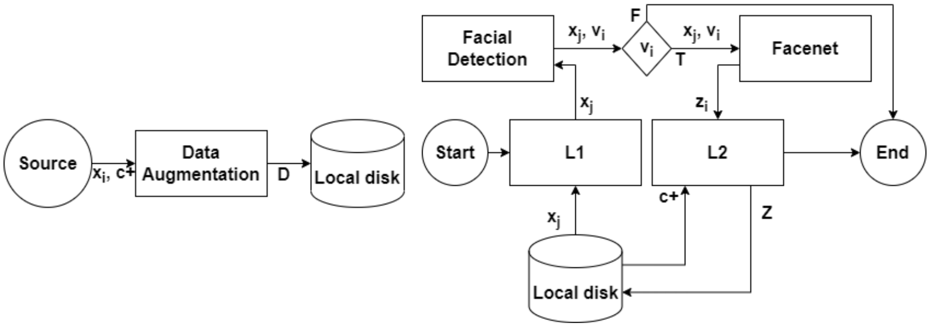
recognition of all persons (users) who show themselves in front of the camera and grant (or not) access to the restricted area.

User registration *User Registration* block registers users that the system will allow access to once they are facially recognized and is divided into two stages that describe the image flow and processing as shown in Figure 2.

As shown in Figure 2-(a), in the first stage, a photograph of the user's face (x_i) is provided frontally by *Source*. Also, enter the user's name ($c+$) by which he/she will be identified.

Before storing the user data (photo and name), two copies of the images are generated in the *Data Augmentation* module. These copies correspond to images derived from the original user's photograph, one with a treatment that brings the user's face closer and the other with a slight change in the angle of the face. In this way, it is sought to cover certain angles or distances of the user in front of the *Camera*. These three images are stored in a directory (D) per user called the user's name. This D is stored in the *Local disk*.

Then, in the second stage, as shown in Figure 2-(b), the *L1* module sends the images extracted from the *Local disk* (x_j) to the *Face Detection* module. *Face Detection* detects



(a) User photo and name capture stage. (b) Feature vector generation and storage.

Figure 2. Extraction of the architecture for the creation of the database.

and extracts faces from an image. This is achieved through models or molds in an XML format called Haar Cascades, which contain features of a frontal face in a numerical data matrix. If one or more regions of the image correspond successfully, these regions are extracted as coordinates within the original image. If an image contains faces, and they are detected, the image and its coordinates (x_j, v_i) are sent to *Facenet* module. This module uses a neural network called FaceNet to extract the characteristics of the detected faces.

CNNs are composed of the feature extraction layer and the learning layer. The feature extraction layer is made up of convolution and reduction (known as pooling) layers, and the learning layer is typically a fully connected neural network. CNNs are designed to process data in multiple arrays, for example, a color image composed of three-dimensional arrays containing pixel intensities in RGB [63].

FaceNet is one of the Deep Learning-based facial recognition applications. It is a one-shot learning algorithm that calculates the similarity distance for each face in Euclidean space [61]. FaceNet [52] was created by Google in 2015, and it is a pre-trained CNN that extracts the main features of faces and transforms them into a 128-dimensional vector known as *embedding*, that represents the principal features extracted from the faces.

Conventional CNNs fail to solve the problem called *one-shot-learning* [12], which is attributed to the inability to do optimal training with only one image per user in the database. In contrast, FaceNet, being a pre-trained CNN with more than 500 million images, according to the authors, achieves a 99.63% effectiveness [52], which is more than the effectiveness reached by the human eye.

FaceNet architecture, as shown in Figure 3, starts from a batch of face thumbnails that works as a Deep Architecture (CNN) input. The Deep architecture transforms the thumbnail image into a 128-D vector. This vector is normalized in L2 process, which results in an Embedding of the face. The *embedding* is followed by Triplet Loss during training.

Facenet module returns to *L2* module a 128-dimensional vector (z_i) corresponding to the features of the face in the image.*L2* also extracts the user's name $(c+)$ and assign its to the z_i. The final result of the images in database treatment is a Pickle file with the extension ".pkl" (Z), which are all the identities of Euclidean positions in a 1x128

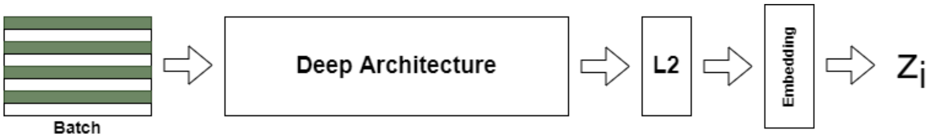


Figure 3. FaceNet structure.

dimensions space stored in the *Local disk* module. Pickle is a Python module used to serialize a file on a disk and then deserialize it back into the program at runtime.

In Z, all the feature vectors of the faces available in D are stored. Each time one or more users are stored, all the feature vectors of the users stored in D must be re-calculated.

User authentication This process requires the interaction of five components as shown in Figure 4. It starts with an image (x_i, 160x160 pixels) sent to the *Face Detection* module. The extracted image and coordinates (x_i, v_i) are dispatched to the *Reality Detection* module.

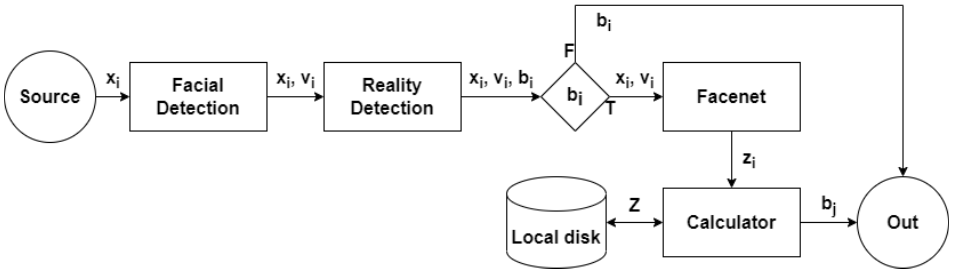


Figure 4. Architecture extraction for facial identification.

Real faces must correspond to real people's faces shown on camera in real-time, while fake faces correspond to attempts to spoof the system using printed or digital photos. In other words, in this module, a classification treatment is carried out with a CNN to the regions obtained from the *Face Detection* module, becoming an anti-spoofing filter for images of non-real people, such as photographs placed in front of the camera.

The *Reality Detection* module uses a typical CNN called ResNet18 that consists of 17 convolution layers and one fully connected network layer, and the activation function used is ReLU [21]. Its purpose is to extract small image regions corresponding to the user's skin parts. Then, remote photoplethysmography (rPPG) treatment detects heart rates and classifies the extracted faces into real or false faces. The rPPG is a heart rate measurement technique that calculates light reflected from the skin, this change is calculated by comparing the reflected light from one frame to the next. The amount of blood in a specific skin area causes capillary dilation and constriction, which causes the light reflected from the skin to change according to the volume of blood circulating [59]. This heart rate measurement method reached a 98.97% test accuracy [56], using a 0.307 M effective pixel camera (ours is 8.08 M effective pixels).

When an image is filtered, and the result is a fake face *Reality Detection* sends a false Boolean (b_i), and the image (x_i) with the face squared in red is displayed on the screen as shown in Figure 7-(c). Otherwise, if the filtered face gives a true result, the *Reality Detection* module sends the original image and its coordinates (x_i, v_i). *Facenet* module is responsible for extracting the features from the faces present in the image and sending a new 128D vector (z_i) to the *Calculator* module. Then, the *Calculator* module calculates the angles from this new vector with all the vectors (Z) extracted from the *Local disk*.

The angles between vectors are calculated by cosine similarity (CS) using the Equation 1 where a_i is the resulting angle, w_i is the feature vector obtained from the new image generated in real-time, and z_i is the representation of the feature vectors stored in *Local disk*. The resulting values range is from 0 to 1 ($a_i \in \mathbb{R}(0, 1)$), where 1 is the biggest angle, and 0 is the smallest. Figure 5 shows an example of a comparison between two vectors, the separation angle and the defined threshold.

$$a_i(w_i, z_i) = 1 - \frac{w_i \cdot z_i}{||w_i||_2 \cdot ||z_i||_2} \tag{1}$$

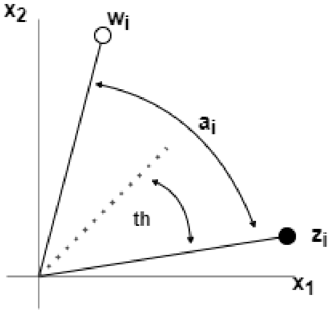


Figure 5. Cosine similarity for comparison of two data vectors.

At this point, if the smallest angle is under a threshold (th), the *Calculator* module extracts the name ($c+$) of the most similar vector, which is a person's name. Otherwise, if the smallest angle is not below the th, the name "Unknown" is assigned. Similarly, if two or more vectors obtain the same CS and are below the th, "Unknown" is assigned.

We implemented a threshold of 0.30 based on the need to ensure greater similarity between the feature vectors of the detected faces, which helps to strengthen the system's security and reduce the risk of confusion between users. We pointed out that decreasing the similarity threshold raises the level of features required for a face to be recognized as valid, resulting in a decrease in False Positive cases.

In contrast, in [53], a th of 0.40 was calculated using the C4.5 [48] algorithm, obtaining a Recall of 96.42, which indicates the presence of False Positive cases. By adjusting the threshold to 0.30 in our work, we seek to mitigate this problem and reduce the probability of misidentifications, prioritizing the security of the system.

It is essential to note that by lowering the threshold, it is possible to increase the rate of False Negatives, i.e., cases in which a legitimate user may be erroneously rejected.

However, in the context of security, it is preferable to go inaccurate in denying access to a legitimate user than to allow access to an intruder. Therefore, by sacrificing some False Negatives in favor of reducing False Positives, we are strengthening the integrity and reliability of the facial recognition system.

Figure 6 shows a simulation of the CS obtained of a user's face taken in real-time (w_i) compared to each of the previously recorded users (z_n). Because z_2 has the smallest angle and is below the th, the output is assigned the z_2 identity.The CS calculation process takes one millisecond (0.001 seconds) on average per registered user, i.e., 1-300 registered users have a minimal impact on the system response time.

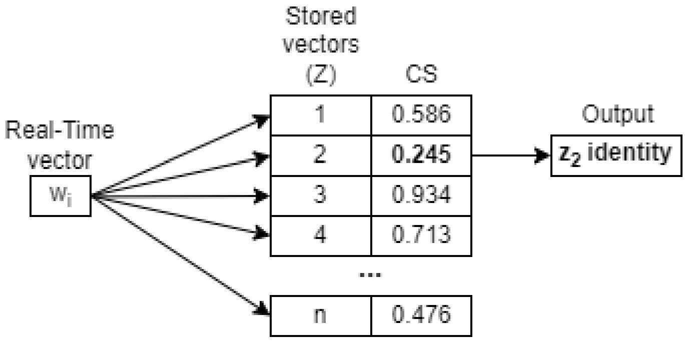
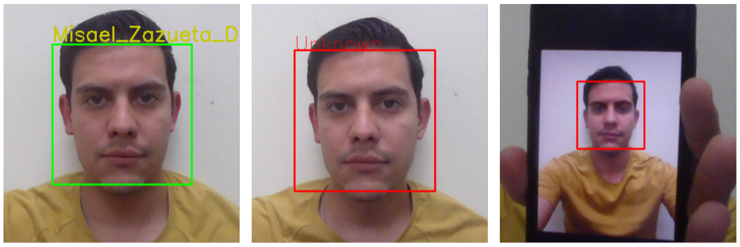


Figure 6. Simulation of angle calculation between vectors for identity assignment.

Finally, the output generated by the *Calculator* module generates a new image created with the same input image (x_i). Still, the detected person's face is framed with a green square with its name on it and displayed on the *Screen* as shown en Figure 7-(a). Also, a Boolean (b_i) is sent to an *Arduino UNO* to unlock the door in case the facial recognition is successful. Otherwise, if the result is unsuccessful, the image of the unrecognized person's face will be displayed on the screen in a red square with the word "Unknown" as shown in Figure 7-(b).



(a) Successful facial recognition in Screen.

(b) Unsuccessful facial recognition in Screen.

(c) Spoof try rejected in Screen.

Figure 7. Possible results showed in Screen.

3.2. User and Control Interfaces

This module is responsible for performing the image acquisition and processing tasks as well as the delivery of results that can be positive or negative for the unlocking or blocking of the door, respectively. The *Jetson Nano* microcomputer, the *Camera*, and the *Screen* are three key system interfaces due to the importance of their tasks. A cabinet was designed and machined to protect and fix these interfaces to contain them in a strategic position for their functions.

The *Camera* acquires images for two elementary purposes in the system. First, it takes photographs to add users to the database used in the *User Registration* process, and second, it captures the users' images that will be facially identified by the *User Authentication* process.

A microcomputer *Jetson Nano* runs the FRS. It receives the images from the *Camera* and sends Boolean signals to an *Arduino UNO* as an output. This device executes the looking and unlocking task sending signals to the *Power Circuit* module. This task is delegated to the *Arduino UNO* to reduce the computational overhead generated on the *Jetson Nano*. The *Screen* receives the output images from the *Jetson Nano*, and it works as visual support so the users can see their faces.

The devices descriptions are:

- *Camera*: IMX219 is a 4.60mm diagonal CMOS (Type 1/ .0) active pixel type image sensor with a square pixel matrix and 8.08M effective pixels. This chip is powered by three power supplies, 2.8V analog, 1.2V digital, and 1.8V IF, and has low power consumption [3].
- *Jetson Nano*: 472 GFLOPS of computing performance with a quad-core 64-bit ARM CPU and an integrated 128-core NVIDIA GPU. It also packs 4GB LPDDR4 memory in an efficient low-power package with 5W / 10W power modes and 5V DC input [39].
- *Screen*: 7 inches touch IPS Screen connected via HDMI, 1024x600 hardware resolution.USB protocol translator, converting the touch signal into standard multi-point touch protocol to achieve smooth multi-point touch control. Powered with 5V via micro-USB.
- *Arduino UNO*: is a development board based on ATmega328P, with 14 terminals for data input or output. It works with a 16 MHz resonator, USB/Power jack connections [4].

3.3. Power Circuit

Since the *Arduino UNO* device is capable of sending electrical signals in the range of 0V to 5V and the *Electronic Lock* is powered with 12V, a *Power Circuit* module is needed to control the current flow required by the *Electronic Lock* with the *Arduino UNO* signals. This *Power Circuit* is composed of three electronic devices: a *Relay*, a *Power Source*, and the aforementioned *Electronic Lock*.

The facial recognition system will send a true or false value depending on the result of the user's identification. When the *Arduino UNO* receives a true value, it sends a 5V signal and activates the *Relay*. This device acts as a “switch”, i.e., when the *Relay* receives the signal, it activates and allows the voltage flow (12V) from the *Power Source* to the *Electronic Lock*. The *Electronic Lock* is a counter-plate that unlocks when receiving a 12V signal. This counter plate is located on the door frame. The door lock prevents the door from opening until the counter plate is energized and allows manipulation.

- *Electronic Lock*: is a fixed device that, upon receiving an electrical signal, its opening mechanism unlocks to allow a door to open.
- *Relay*: is an electronic device that distributes an electrical signal in two directions (A or B). When the relay is in a normal state, it sends the signal in direction A, but when it is activated, it sends the signal in direction B.
- *Power Source*: is an electronic device that transforms electrical energy from alternating current into direct current. It is commonly used to transform ordinary household electrical energy (120 volts AC) into electrical energy for charging mobile devices (5/12/24 volts direct current).

4. Testing and results

The Facial Recognition System (FRS) was developed, compiled, and executed in Python language. It uses Tensorflow-GPU and Keras libraries to implement CNNs for reality detection and facial recognition (ResNet18 and FaceNet). In the same way, the “switch” process executed by the Arduino UNO was developed, compiled, and executed in C++.

Three stages of FRS testing were performed: (i) face recognition testing with artificial lighting, (ii) face recognition testing with natural lighting, and finally, (iii) reality detection testing using printed and digital photographs and videos.

The individuals considered for testing the system represent a mixed sample of Mexican people aged 14 to 40. Approximately 90% of the sample is in the 18 to 23 age range. The skin tones of the individuals are diverse, ranging from tan to lighter skin tones.

The first testing stage was performed with artificial lighting on 85 users, 72 registered users, and 13 non-registered users. The tests were conducted with illumination on the top of the users’ heads to avoid shadows that would hinder facial recognition. The population of test users registered 425 trials (5 for each user) to the FRS. From 360 registered-users trials, the system successfully recognized 321 times (TP, True Positive) and 39 times failed to identify the person (FN, False Negative), giving the result of “Unknown". No identity assignment errors were recorded, i.e., the system did not confuse any user (registered or not) with another (FP, False Positive). Within the controlled environment testing stage, 13 users who were not registered in the system were tested. The total number of attempts was 65. 100% of the cases were identified as “U", i.e., of all the cases predicted as “Unknown", the real result the system gave was the same (TN, True Negative). Table 1 shows the confusion matrix resulting from the tests in a controlled environment.

Tables 2 to 6 show the relationship matrices related to face recognition performed in a controlled environment on 72 users (S) registered in the system. User number 23 was recognized only once by the FRS, and 4 obtained a “U" (Unknown) result. During the execution of the tests, different variables were observed that influenced the facial identification of some test subjects, e.g., face pose, glasses, or fringe. If any of these variables

Table 1. Resulting confusion matrix for the controlled environment testing stage.

		Predicted	
		Positive	Negative
Real	Positive	321	0
	Negative	39	65

changed from the initial sampling, facial recognition showed delays and failed identifications. Therefore, as users attempted facial recognition and failed to make a successful identification, they were asked to change these variables to cover the issues surrounding the face presentation to the system.

Table 2. Relationship matrix of face recognition results with artificial lighting to registered users (1/5)

		Predicted															
		S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**
Real	**S1**	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	S2	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	S3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	S4	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
	S5	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0
	S6	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0
	S7	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0
	S8	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0
	S9	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0
	S10	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0
	S11	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0
	S12	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0
	S13	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0
	S14	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0
	S15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0
	S16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	“U"	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	3

Table 3. Relationship matrix of face recognition results with artificial lighting to registered users (2/5)

		Predicted															
		S17	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**	**S25**	**S26**	**S27**	**S28**	**S29**	**S30**	**S31**	**S32**
Real	**S17**	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	S18	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	S19	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0
	S20	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
	S21	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0
	S22	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0
	S23	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
	S24	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0
	S25	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0
	S26	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0
	S27	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0
	S28	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0
	S29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	S30	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0
	S31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0
	S32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
	“U"	0	0	0	0	0	0	4	0	0	0	0	0	5	0	0	0

Table 4. Relationship matrix of face recognition results with artificial lighting to registered users (3/5)

		Predicted															
		S33	**S34**	**S35**	**S36**	**S37**	**S38**	**S39**	**S40**	**S41**	**S42**	**S43**	**S44**	**S45**	**S46**	**S47**	**S48**
Real	**S33**	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	S34	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	S35	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0
	S36	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
	S37	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0
	S38	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0
	S39	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0
	S40	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0
	S41	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
	S42	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
	S43	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0
	S44	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0
	S45	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0
	S46	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0
	S47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0
	S48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
	“U"	0	0	0	0	1	0	0	0	4	4	0	0	0	0	0	0

Table 5. Relationship matrix of face recognition results with artificial lighting to registered users (4/5)

		Predicted															
		S49	**S50**	**S51**	**S52**	**S53**	**S54**	**S55**	**S56**	**S57**	**S58**	**S59**	**S60**	**S61**	**S62**	**S63**	**S64**
Real	**S49**	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	S50	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	S51	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0
	S52	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
	S53	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0
	S54	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0
	S55	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0
	S56	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0
	S57	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0
	S58	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0
	S59	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0
	S60	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0
	S61	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0
	S62	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0
	S63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0
	S64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
	“U"	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0

Table 6. Relationship matrix of face recognition results with artificial lighting to registered users (5/5)

		Predicted							
		S65	**S66**	**S67**	**S68**	**S69**	**S70**	**S71**	**S72**
Real	**S65**	0	0	0	0	0	0	0	0
	S66	0	5	0	0	0	0	0	0
	S67	0	0	5	0	0	0	0	0
	S68	0	0	0	4	0	0	0	0
	S69	0	0	0	0	5	0	0	0
	S70	0	0	0	0	0	5	0	0
	S71	0	0	0	0	0	0	0	0
	S72	0	0	0	0	0	0	0	5
	“U"	5	0	0	1	0	0	5	0

The faces of the users with the highest number of failed attempts showed a trend in the images captured during the registration stage. For example, the presence of severe occlusions in the eye area is notorious because the users had very thick eyebrows, wore glasses and/or wore very pronounced bangs. Another factor detected in the registration images is the darkening of the mouth area due to pronounced commissures. These factors produce occlusions and shadows in critical areas of the user's face during registration, thus losing essential features of the user's face and making it difficult to authenticate users during testing is shown in Tables 2 to 6. Similarly, make-up influences the recognition of some users, i.e., if the user presents different make-up conditions during registration and authentication, the BLS will tend to fail due to changes in facial features.

The second testing stage was performed outdoors under natural light conditions during sunset. Table 7 shows the results obtained from testing 10 test users registered in the system between 19:30 and 19:45 GMT-6. It is a time when the sun begins to set, and the amount of light decreases considerably but without reaching total darkness. Testing outdoors is to add a variable related to the amount of light available for face recognition and reality detection.

At the end of this stage, the system failed to recognize 6 out of 50 attempts. Also, within this stage, 25 tests were performed on five subjects not registered in the system. As in the previous stage, the system did not recognize unregistered subjects. Table 7 shows the 44 facial recognition hits (TP), 25 successes in non-face recognition (TN), six unsuccessful recognitions (FN), and 0 identity assignment errors (FP) obtained.

Table 7. Resulting confusion matrix for the uncontrolled environment testing stage

		Predicted	
		Positive	Negative
Real	Positive	44	0
	Negative	6	25

Table 8 shows the relationship matrix of the tests performed on users with natural lighting.

Table 8. Relationship matrix of face recognition results with natural lightning to registered subjects

		Predicted									
		S73	**S74**	**S75**	**S76**	**S77**	**S78**	**S79**	**S80**	**S81**	**S82**
Real	**S73**	5	0	0	0	0	0	0	0	0	0
	S74	0	5	0	0	0	0	0	0	0	0
	S75	0	0	4	0	0	0	0	0	0	0
	S76	0	0	0	5	0	0	0	0	0	0
	S77	0	0	0	0	5	0	0	0	0	0
	S78	0	0	0	0	0	2	0	0	0	0
	S79	0	0	0	0	0	0	5	0	0	0
	S80	0	0	0	0	0	0	0	3	0	0
	S81	0	0	0	0	0	0	0	0	5	0
	S82	0	0	0	0	0	0	0	0	0	5
	“U"	0	0	1	0	0	3	0	2	0	0

As in the previous testing stage, the users who presented more errors in the testing stage, as shown in Table 8, demonstrated more shadows and occlusions on the face in the capture of log images.

As the final results of the facial recognition tests in the first two stages, Table 9, 365 hits in face recognition (TP), 45 system failures in confusing registered subjects with unknown subjects (FN), 90 hits in non-recognition of non-registered subjects (TN), and 0 mistaken identity assignments to registered or non-registered subjects (FP) can be observed.

Table 9. Resulting confusion matrix from the two face recognition stages of testing

		Predicted	
		Positive	Negative
Real	Positive	365	0
	Negative	45	90

Once the two stages of face recognition testing were completed, results were obtained for the evaluation criteria: false rejection rate (FRR), false acceptance rate (FAR), accuracy, and F1. FRR (Equation 2) is a binary classification measure that provides an evaluation of how well the system avoids false negatives (FN). FAR (Equation 3) evaluates how well false positives are avoided. Both measures generate values between 0 and 1; the closer the result is to 0, the better the system evaluation. Accuracy is the most widely used measurement in classification systems. It is calculated by dividing the total number of hits by the total number of samples, i.e., accuracy is the percentage of hits in the system (Equation 4).

$$FRR \text{ (False rejection rate)} = 1 - TP/(TP + FN) \tag{2}$$

$$FAR \text{ (False acceptation rate)} = 1 - TN/(TN + FP) \tag{3}$$

$$Accuracy = (TP + TN)/(TP + FP + FN + TN) \tag{4}$$

The evaluation criteria described above were applied to the results corresponding to the face recognition tests in the two stages shown in Tables 1, 7, and 9. The evaluations obtained are summarized in Table 10.

Table 10. FRS evaluation criteria results of tests in the face recognition stages

		Evaluation criteria		
		FRR	FAR	Accuracy
Conditions	Artificial lightning	0.1083	0.0	0.9082
	Natural lightning	0.1200	0.0	0.9200
	Average	0.1098	0.0	0.9100

To verify that the system is robust to different types of lighting (natural or artificial), a paired T-Student [25] distribution parametric analysis was performed. We proposed the next hypotheses:

- H_0: *No significant difference exists in tests performed with natural and artificial illumination.*
- H_1: *There is a significant difference in tests performed with natural and artificial illumination.*

A random sampling of 10 users from the artificial lighting test stage was generated for comparison. Using the individual effectiveness obtained for each user (20% per successful recognition) showed in Tables 2 to 6 and Table 8, the following data were obtained: mean difference $\bar{x} = -6$, sample size $n = 10$, standard deviation $S_d = 32.73$. The degrees of freedom ($d_f = 9$) and significance level ($\alpha = 0.05$) are considered for the parametric analysis. The corresponding T-Student statistic value is $t_\alpha : 0.005, df : 9 = 2.2622$. Equation 5 was used and $t = -0.5797$ was obtained.

$$t = \frac{\bar{x}}{\frac{S_d}{\sqrt{n}}} = \frac{-6}{\frac{32.73}{\sqrt{9}}} = -0.5797 \quad (5)$$

Figure 8 shows that the t-value (statistic) obtained is within the confidence interval. H_0 is accepted, i.e., no significant difference exists between the tests performed with artificial and natural lighting. Because of this, the system is robust to changes in illumination from natural to artificial or vice versa.

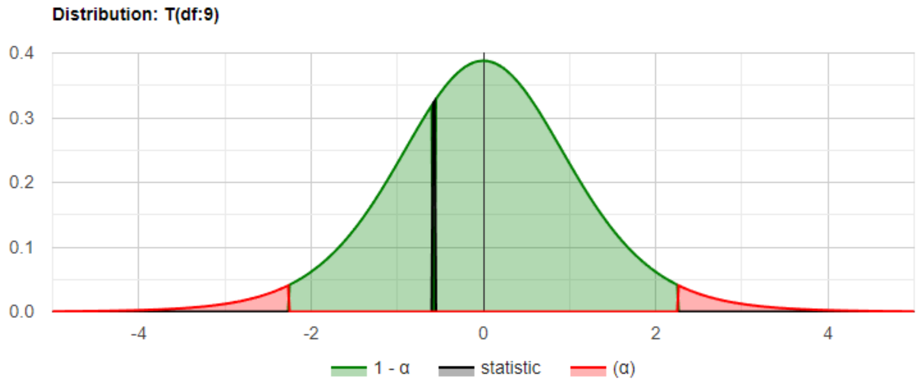


Figure 8. Paired T-Student distribution parametric analysis graph

In the last testing stage, 90 photographs and ten videos were submitted to the FRS. None of the attempts were successful in spoofing the system (TN). In addition, the system was recognized as real people in all 500 attempts in the previous two stages of testing (TP). Thus, as shown in Table 11, the *Reality Detection* module obtained 500 TP and 100 TN.

Table 11. Resulting confusion matrix for the reality detection testing stage

		Predicted	
		Positive	Negative
Real	Positive	500	0
	Negative	0	100

4.1. Discussion

In Table 12, we compared different works considering hardware, techniques involved for facial recognition, and the performance results obtained.

For the hardware comparison, we show that the control interfaces of [15,26,41,51,60] are based on different Raspberry Pi models. Some advantages of this device are its low cost, its small size, which makes it easy to hide inside the room and its ease of use. On the other hand, [35] and our proposal are based on Jetson Nano, which, unlike any Raspberry Pi model, has an integrated GPU that accelerates the execution of algorithms. However, it is larger in size and higher in cost. The devices used as power circuitry vary in all cases but demonstrate a similarity of operation in power management. Hardware costs vary in two price ranges, from $45 to $60, and from $225 to $250 USD. This is due to the incorporation of a Neural Compute Stick 2 ([41]) and the integrated GPU in the Jetson Nano ([35], ours), which provide additional support for image data processing and improved response speed.

In [35,41,60] and our proposal, pre-trained One-Shot-Learning MobileNet and FaceNet networks were used for face recognition. The main advantage of implementing this type of network is saving training time that is usually required to properly input a neural network. In addition, FaceNet was trained with more than 500 million images, while MobileNet with 3 million. In the remaining cases, they trained their own models for facial identification, and being less dense, they require less computational power and generally less execution time. While in the proposals [15,26,51], it is necessary to retrain the facial identification models each time a new user is required to register. This is because the models need to learn to classify each of the classes created.

However, [26, 35, 60] exceed 96.00% in terms of successful recognitions. It is necessary to remark that these proposals do not have an anti-spoofing stage, which is totally inconvenient for face recognition tasks in security despite having a good response time ([35]). [15] shows an effectiveness rate of 92.68%, but due to its anti-spoofing method, it requires more than 10 seconds to open a door and an additional action (blink) by the user. Our proposal is based on a CNN ResNet18 that makes image extractions corresponding to parts of the skin and an rPPG algorithm that calculates the person's heart rate by the change of illumination on the skin caused by blood flow. This means that a printed or digital image will not cause such illumination changes in the skin.

Most of the work tested in artificially lit areas, only [26] demonstrated an average effectiveness of 97.5% in daylight conditions at four different times. [35] demonstrated an effectiveness of 97.7% and a response time of 0.47 s, making it competitive for face recognition tasks that require high response speed and do not require an anti-spoofing system. The FAR demonstrated in [41] was 35.71%, which is not suitable for face recognition systems. It means the system allowed access to 35 unauthorized people out of 100.

Our proposal demonstrated an average effectiveness of 91% and a response speed of 1.68s, which competes with the related work described above. The resulting average FRR was 1.098%, corresponding to the 45 false-negative results, which denied access to users registered in the system. In addition, a FAR of 0% represents non-existent false positives. This shows that the system errors consist solely of false rejections. Of these, 95.12% of the subjects succeeded in gaining access on at least one attempt. The system recognized most of the users who were initially not recognized in their first attempt, after a small adjustment of the face position with respect to the camera to reduce shadows, center the

Table 12. Comparison of hardware, software, and performance of related work and the proposal

		References							
		[41]	[35]	[26]	[60]	[15]	[51]		**Proposed**
Hardware	Procesor	Raspberry + NCS2 $245	Jetson Nano $225	Raspberry Pi $45	Raspberry Pi 4 $61	Raspberry Pi 3 B+ $50	Raspberry Pi $45		Jetson Nano $225
	Power circuit	Coil + Electro-magnetic Lock	Reed Switch	Relay + Solenoid	Servo-motor	Relay + Solenoid	MOSFET + Solneoid		Relay + Electronic Lock + Power Source
Software	Recognition	MobileNet	FaceNet	CNN	ResNet + FaceNet	HOG + SVM	LBPH	HOG	FaceNet
	Liveness	SSD	-	-	-	Blink detection	-	-	ResNet18 + rPPG
Performance	Recognition tests	80	1260	120	230	-	200	200	500
	Reality tests	5	-	-	-	-	-	-	600
	Recognition accuracy	88.75%	97.70%	Morning = 96.67% Afternoon = 96.67% Evening = 96.67% Night = 100% AVG = 97.5%	87.36%	92.68%	98.00%	96.00%	CE = 90.82% UE = 92.00% AVG = 91.00%
	FRR	0%	-	0%	-	-	4%	2%	CE = 1.083% UE = 1.2% AVG = 1.098%
	FAR	35.71%	-	Morning = 6.25% Afternoon = 6.25% Evening = 6.25% Night = 0% AVG = 4.76%	-	-	0%	6%	0%
	Reality accuracy	100%	-	-	-	-	-	-	100%
	Execution	-	0.47 s	-	-	10.196 s	-	-	1.68 s

face in the image, among other aspects that could affect the recognition. But, we remark that it does not grant access to unregistered users and does not confuse one user with another. In this work, numerous tests on both aspects of user authentication (recognition and reality) demonstrate that the system shows a balance of effectiveness, FAR, FRR, and speed of response.

Table 13. Top selled Amazon facial recognition based door locks

Model	Price (USD)
Necchuizo	$365
We Technology	$249
NGP	$399

The Table 13 shows the top three best-selling commercial door locking devices with facial recognition, according to Amazon.com in the US. They consist of electronic devices that completely replace the door lock in question. In none of the cases are technical specifications of effectiveness, FAR, or FRR given. To operate, they require batteries every so often. Being all-in-one devices, they are challenging to maintain in case any of the components fail. Likewise, they require mobile applications to control their user registration and deletion operations, including WiFi connectivity.

Modulating the hardware provides high maintainability to the BLS because each device works individually and can be easily repaired or replaced. Also, this modulation facilitates installation or change of location. On the other hand, there are similar commercial locks on the market that achieve better aesthetics and smaller size, but because of this, they increase the difficulty of maintenance and location changes. In addition, these locks operate with disposable or rechargeable batteries, which is a disadvantage if replacement is omitted in a timely manner. Finally, most commercial locks have a registration limit of 100 users, while FaceNet was trained with 10000 identities.

5. Conclusions

This paper presents a proposal for a biometric-lock system (BLS) that works as a facial recognition lock. This system limits access to restricted areas by unlocking doors to users previously registered as authorized persons.

We divide the development of this BLS into three main blocks. The first one refers to the user and control interfaces, which are the devices used to interact with the user and perform facial recognition tasks. The second block is composed of three devices, a relay that controls the voltage flow, a power supply that generates the required voltage, and an electronic lock that locks or unlocks the door opening. Finally, the third block describes the development of the software for the registration and authentication of the users that will interact with the system.

The biggest challenge during the development of the BLS was to produce facial recognition software that is robust to photo and/or video attacks without the use of specialized hardware and that is capable of running on embedded devices. Identifying real people from non-real people is an extremely important feature as it represents a significant security aspect of facial recognition systems. Also, it is important to mention that there are some works in the state of the art that achieve these objectives using different techniques (blink detection, human detection). However, these techniques require actions by the user, increasing their difficulty of use, and this causes a considerable increase in response time. Because our proposal measures the heart rate at the same image of the user's face, it decreases user cooperation and response time.

Tests performed on the BLS show an effectiveness of 90.82% in artificial light and 92.00% in natural light. A T-Student test was performed to verify if there was a significant difference between the two lighting conditions. The results indicate the system's resilience to changes in illumination, showcasing no significant disparities. On average, it achieved an effectiveness rate of 91.00% across all facial recognition trials, with the remaining 9% representing instances where user identification faltered. Notably, the system exhibited a flawless False Acceptance Rate (FAR) of 0%, ensuring zero false positives or user confusion. Furthermore, in a notable statistic, 78 out of 82 users (95.12%) successfully gained access after at least one of five attempts, demonstrating the system's ability to recognize users even if initial recognition fails by adjusting the facial positioning relative to the camera. Also, a 100% effectiveness rate was obtained in the anti-attack tests on the system regarding the test carried out in the experiment. Finally, the average response time obtained was 1.68 seconds. Nevertheless, an average FRR of 1.098% was obtained since some users presented occlusions in important face areas, such as the eyes and mouth. These factors represent a future challenge to reduce the FRR and grant access to registered users.

The results showed the competitiveness of the proposed BLS compared to the works presented in the state of the art. However, it does not cover the security scenarios required by maximum security buildings; banks, vaults, or prisons. In other words, the BLS has the necessary effectiveness and speed to be implemented in medium security restricted access locations such as offices, classrooms, houses, among others. The commercial devices shown in Table 13 are generally put to the same use. Although these devices currently on the market have significantly decreased in price, the modularization of our proposal could considerably improve the cost-benefit of operating two or more locks with the same Jetson Nano.

In future work, we will focus on improving the effectiveness of facial recognition by implementing different image resolutions, modifying the thresholds for facial identification, or using other CNNs for image processing, such as MobileNet, VGGFace, and DeepFace. Also, different actuators for door opening will be implemented to cover more applications; magnetic locks, servo-motors, and pneumatic or hydraulic pistons, among others. Finally, increase the number of doors operated by the same device to measure operability.

References

1. Ali, W., Tian, W., Din, S.U., Iradukunda, D., Khan, A.A.: Classical and modern face recognition approaches: a complete review. Multimedia tools and applications 80, 4825–4880 (2021)
2. Arashloo, S.R., Kittler, J.: Efficient processing of mrfs for unconstrained-pose face recognition. In: 2013 IEEE sixth international conference on biometrics: theory, applications and systems (BTAS). pp. 1–8. IEEE (2013)
3. ArduCam: Camera sony imx219, https://www.arducam.com/product/arducam-raspberry-pi-camera-v2-8mp-ixm219-b0103/, web accessed in: 07-06-2022
4. Arduino: Arduino uno r3, https://docs.arduino.cc/hardware/uno-rev3, web accessed in: 07-06-2022
5. Augusto, J.C., McCullagh, P.: Ambient intelligence: Concepts and applications. Computer Science and Information Systems 4(1), 1–27 (2007)
6. Bhatt, H.S., Bharadwaj, S., Singh, R., Vatsa, M.: Recognizing surgically altered face images using multiobjective evolutionary algorithm. IEEE Transactions on Information Forensics and Security 8(1), 89–100 (2012)
7. Bud, A.: Facing the future: The impact of apple faceid. Biometric technology today 2018(1), 5–7 (2018)
8. Chen, S., Shen, J., You, X., Chen, J., Yu, C.: A dynamic cryptography door lock system based on visible light communication. In: 2018 23rd Opto-Electronics and Communications Conference (OECC). pp. 1–2 (2018)
9. Cui, Z., Li, W., Xu, D., Shan, S., Chen, X.: Fusing robust face region descriptors via multiple metric learning for face recognition in the wild. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3554–3561 (2013)
10. Das, P.K., Hu, B., Liu, C., Cui, K., Ranjan, P., Xiong, G.: A new approach for face anti-spoofing using handcrafted and deep network features. In: 2019 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI). pp. 33–38. IEEE (2019)

11. Ding, C., Tao, D.: Robust face recognition via multimodal deep face representation. IEEE transactions on Multimedia 17(11), 2049–2058 (2015)
12. Eloff, R., Engelbrecht, H.A., Kamper, H.: Multimodal one-shot learning of speech and images. In: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 8623–8627 (2019)
13. Gałka, J., Masior, M., Salasa, M.: Voice authentication embedded solution for secured access control. IEEE Transactions on Consumer Electronics 60(4), 653–661 (2014)
14. Ganjoo, R., Purohit, A.: Anti-spoofing door lock using face recognition and blink detection. In: 2021 6th International Conference on Inventive Computation Technologies (ICICT). pp. 1090–1096. IEEE (2021)
15. Ganjoo, R., Purohit, A.: Anti-spoofing door lock using face recognition and blink detection. In: 2021 6th International Conference on Inventive Computation Technologies (ICICT). pp. 1090–1096 (2021)
16. Ghorbani, M., Targhi, A.T., Dehshibi, M.M.: Hog and lbp: Towards a robust face recognition system. In: 2015 Tenth International Conference on Digital Information Management (ICDIM). pp. 138–141 (2015)
17. González-Jiménez, D., Alba-Castro, J.L.: Toward pose-invariant 2-d face recognition through point distribution models and facial symmetry. IEEE Transactions on Information Forensics and Security 2(3), 413–429 (2007)
18. Goud, K.N., Sindhuri, K.: Enhanced Security for Smart Door Using Biometrics and OTP, pp. 517–526. Springer International Publishing, Cham (2022), https://doi.org/10.1007/978-3-030-96634-8_47
19. Grassi, P.A., Garcia, M.E., Fenton, J.F.: NIST special publication 800-63b: Digital identity guidelines. Tech. Rep. 800-63B, National Institute of Standards and Technology (June 2017), https://doi.org/10.6028/NIST.SP.800-63B
20. Hafez, S.F., Selim, M.M., Zayed, H.H.: 2d face recognition system based on selected gabor filters and linear discriminant analysis lda. arXiv preprint arXiv:1503.03741 (2015)
21. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
22. Hearst, M., Dumais, S., Osuna, E., Platt, J., Scholkopf, B.: Support vector machines. IEEE Intelligent Systems and their Applications 13(4), 18–28 (1998)
23. Hemalatha, A., Gandhimathi, G.: Rfid, password and otp based door lock system using 8051 microcontroller. International Journal of Engineering Research and Technology 7(11), 1–6 (2019)
24. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: Efficient convolutional neural networks for mobile vision applications (2017), https://arxiv.org/abs/1704.04861
25. Hsu, H., Lachenbruch, P.A.: Paired t test. Wiley StatsRef: statistics reference online (2014)
26. Irjanto, N.S., Surantha, N.: Home security system with face recognition based on convolutional neural network. (IJACSA) International Journal of Advanced Computer Science and Applications 11(11) (2020)
27. Jayamaha, R.G.M.M., Senadheera, M.R.R., Gamage, T.N.C., Weerasekara, K.D.P.B., Dissanayaka, G.A., Kodagoda, G.N.: Voizlock - human voice authentication system using hidden markov model. In: 2008 4th International Conference on Information and Automation for Sustainability. pp. 330–335 (2008)
28. Jovanović, B., Milenković, I., Bogićević Sretenović, M., Simić, D.: Extending identity management system with multimodal biometric authentication. Computer Science and Information Systems/ComSIS 13(2), 313–334 (2016)
29. Khowaja, S.A., Lee, S.L.: Hybrid and hierarchical fusion networks: a deep cross-modal learning architecture for action recognition. Neural Computing and Applications 32(14), 10423–10434 (2020)

30. Khowaja, S.A., Lee, S.L.: Semantic image networks for human action recognition. International Journal of Computer Vision 128(2), 393–419 (2020)
31. Komol, M.M.R., Podder, A.K., Ali, M.N., Ansary, S.M.: Rfid and finger print based dual security system: A robust secured control to access through door lock operation. American Journal of Embedded Systems and Applications 6(1), 15–22 (2018)
32. Kortli, Y., Jridi, M., Al Falou, A., Atri, M.: Face recognition systems: A survey. Sensors 20(2) (2020), https://www.mdpi.com/1424-8220/20/2/342
33. Kumar, A., Ravikanth, C.: Personal authentication using finger knuckle surface. IEEE Transactions on Information Forensics and Security 4(1), 98–110 (2009)
34. Lewandowska, M., Nowak, J.: Measuring pulse rate with a webcam. Journal of Medical Imaging and Health Informatics 2(1), 87–92 (2012)
35. Lindner, T., Wyrwal, D., Bialek, M., Nowak, P.: Face recognition system based on a single-board computer. In: 2020 International Conference Mechatronic Systems and Materials (MSM). pp. 1–6 (2020)
36. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C.: Ssd: Single shot multibox detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) Computer Vision – ECCV 2016. pp. 21–37. Springer International Publishing, Cham (2016)
37. Mian, A., Bennamoun, M., Owens, R.: An efficient multimodal 2d-3d hybrid approach to automatic face recognition. IEEE transactions on pattern analysis and machine intelligence 29(11), 1927–1943 (2007)
38. Nortje, L.: Direct and indirect multimodal few-shot learning of speech and images. Ph.D. thesis, Stellenbosch: Stellenbosch University (2020)
39. NVIDIA: Jetson nano developer kit, https://developer.nvidia.com/embedded/jetson-nano-developer-kit, web; accessed in 01-29-2021
40. Organization, U.N.: Unodc burglary 2018. https://dataunodc.un.org/data/crime/burglary, accessed: 2022-03-23
41. Orna, G., Benitez, D.S., Perez, N.: A low-cost embedded facial recognition system for door access control using deep learning. In: 2020 IEEE ANDESCON. pp. 1–6 (2020)
42. Pacheco, J., Miranda, K.: Design of a low-cost nfc door lock for a smart home system. In: 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). pp. 1–5 (2020)
43. Patil, B., Mahajan, V., Suryawanshi, S., Pawar, M.: Automatic door lock system using pin on android phone. International Research Journal of Engineering and Technology (IRJET) 5(11), 1007–1011 (2018)
44. Pillai, J.K., Patel, V.M., Chellappa, R., Ratha, N.K.: Secure and robust iris recognition using random projections and sparse representations. IEEE transactions on pattern analysis and machine intelligence 33(9), 1877–1893 (2011)
45. Prasad, S., Govindan, V., Sathidevi, P.: Palmprint authentication using fusion of wavelet and contourlet features. Security and Communication Networks 4(5), 577–590 (2011)
46. Prity, S.A., Afrose, J., Hasan, M.: Rfid based smart door lock security system. American Journal of Sciences and Engineering Research E-ISSN-2348-703X 4(3) (2021)
47. Queirolo, C.C., Silva, L., Bellon, O.R., Segundo, M.P.: 3d face recognition using simulated annealing and the surface interpenetration measure. IEEE transactions on pattern analysis and machine intelligence 32(2), 206–219 (2009)
48. Quinlan, J.R.: Improved use of continuous attributes in c4. 5. Journal of artificial intelligence research 4, 77–90 (1996)
49. Ranjan, R., Sankaranarayanan, S., Castillo, C.D., Chellappa, R.: An all-in-one convolutional neural network for face analysis. In: 2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017). pp. 17–24 (2017)
50. Ren, J., Jiang, X., Yuan, J.: Relaxed local ternary pattern for face recognition. In: 2013 IEEE international conference on image processing. pp. 3680–3684. IEEE (2013)

51. Saputra, R., Surantha, N.: Smart and real-time door lock system for an elderly user based on face recognition. Bulletin of Electrical Engineering and Informatics 10(3), 1345–1355 (2021), `https://beei.org/index.php/EEI/article/view/2955`
52. Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: A unified embedding for face recognition and clustering. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2015), `http://dx.doi.org/10.1109/CVPR.2015.7298682`
53. Serengil, S.I., Ozpinar, A.: Lightface: A hybrid deep face recognition framework. In: 2020 Innovations in Intelligent Systems and Applications Conference (ASYU). pp. 1–5 (2020)
54. Sidiropoulos, G.K., Kiratsa, P., Chatzipetrou, P., Papakostas, G.A.: Feature extraction for finger-vein-based identity recognition. Journal of Imaging 7(5), 89 (2021)
55. Song, K.C., Yan, Y.H., Chen, W.H., Zhang, X.: Research and perspective on local binary pattern. Acta Automatica Sinica 39(6), 730–744 (2013), `https://www.sciencedirect.com/science/article/pii/S1874102913600518`
56. Tang, C., Lu, J., Liu, J.: Non-contact heart rate monitoring by combining convolutional neural network skin detection and remote photoplethysmography via a low-cost camera. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops (June 2018)
57. Thavalengal, S., Andorko, I., Drimbarean, A., Bigioi, P., Corcoran, P.: Proof-of-concept and evaluation of a dual function visible/nir camera for iris authentication in smartphones. IEEE Transactions on Consumer Electronics 61(2), 137–143 (2015)
58. Thavalengal, S., Bigioi, P., Corcoran, P.: Iris authentication in handheld devices-considerations for constraint-free acquisition. IEEE Transactions on Consumer Electronics 61(2), 245–253 (2015)
59. Wang, W., Den Brinker, A.C., Stuijk, S., De Haan, G.: Algorithmic principles of remote ppg. IEEE Transactions on Biomedical Engineering 64(7), 1479–1491 (2016)
60. Waseem, M., Khowaja, S.A., Ayyasamy, R.K., Bashir, F.: Face recognition for smart door lock system using hierarchical network. In: 2020 International Conference on Computational Intelligence (ICCI). pp. 51–56 (2020)
61. William, I., Ignatius Moses Setiadi, D.R., Rachmawanto, E.H., Santoso, H.A., Sari, C.A.: Face recognition using facenet (survey, performance test, and comparison). In: 2019 Fourth International Conference on Informatics and Computing (ICIC). pp. 1–6 (2019)
62. Wimmer, G., Prommegger, B., Uhl, A.: Finger vein recognition and intra-subject similarity evaluation of finger veins using the cnn triplet loss. In: 2020 25th International Conference on Pattern Recognition (ICPR). pp. 400–406 (2021)
63. Wolf, L., Hassner, T., Maoz, I.: Face recognition in unconstrained videos with matched background similarity. In: CVPR 2011. pp. 529–534. IEEE (2011)
64. Xi, S., Yang, L., Zhao, Y., et al.: A practical design for face recognition with anti-spoofing based on non-visible light cameras. Academic Journal of Computing & Information Science 3(2) (2020)
65. Yan, Z., Zhao, S.: A usable authentication system based on personal voice challenge. In: 2016 International Conference on Advanced Cloud and Big Data (CBD). pp. 194–199. IEEE (2016)
66. Yang, W., Hu, J., Wang, S.: A delaunay quadrangle-based fingerprint authentication system with template protection using topology code for local registration and security enhancement. IEEE transactions on Information Forensics and Security 9(7), 1179–1192 (2014)

José Misael Burruel Zazueta received a degree in mechatronics engineering in 2018 and a master's degree in computer science in 2021, both from the Tecnológico Nacional de México campus Culiacan. During his master's degree he developed research works in face recognition based on deep learning, generating research articles accepted in prestigious conferences in Mexico. Also, he has collaborated in research related to wind time series forecasting with a paper published in IEEE. He has supervised four engineering thesis. He is currently pursuing a PhD in engineering sciences, focusing on biometric recognition based on artificial intelligence techniques.

Héctor Rodríguez Rangel is a graduate of the Technological Institute of Morelia, obtaining a Computer Systems bachelor's degree (2009). The master's and doctorate degrees were carried out in the Postgraduate Department of the Faculty of Electrical Engineering at UMSNH (2009, 2014). He is currently a Full-time Professor at the Technological Institute of Culiacán a and a member of the Mexican National System of Researchers, distinguished as Level 1. During his doctoral degree, he made an academic stay at the University of Oregon. In his doctorate, he worked on a qualitative bifurcation diagram project. His lines of research focus on Optimization, Intelligent Computing, and Pattern Recognition.

Gloria Ekaterine Peralta-Peñuñuri is a full-time teacher in the Technological Institute of Culiacán. In recent years, she has participated in financed research projects together with researchers from the Technological Institute of Culiacán, where several postgraduate and bachelor's degree students have participated. Gloria is currently in the second semester of her Ph.D. program. Her lines of research focus on Education and Applied Intelligent Computing.

Vicenç Puig received the telecommunications engineering degree in 1993 and the Ph.D. degree in Automatic Control, Vision, and Robotics in 1999, both from Universitat Politècnica de Catalunya (UPC). He is Full Professor of the Automatic Control Department and a Researcher at the Institut de Robòtica i Informàtica Industrial, both from the UPC. He is currently the Head of the Research Group in Advanced Control Systems and the Responsible of the PhD Programme in Automatic Control and Robotics. at UPC. Formerly, he was the Director of the Automatic Control Department. He has developed important scientific contributions in the areas of fault diagnosis and fault tolerant control using interval and linear-parameter-varying models using set-based approaches. He has participated/leaded more than 20 European and national research projects in the last decade. He has also led many private contracts with several companies and has published more than 120 journal articles and more than 450 in international conference/ workshop proceedings. He has supervised over 20 Ph.D. dissertations and over 40 master's theses/final projects. He is currently the chair of the IFAC Safeprocess TC Committee 6.4 (2020-until now) and was the vice chair (2014–2017). He has been the general chair of the Third IEEE Conference on Control and Fault-Tolerant Systems (Systol 2016 and 2021) and the IPC chair of the IFAC Safeprocess 2018.

Ignacio Algredo-Badillo received the B.Eng in Electronic Engineering from Technologic Institute of Puebla (ITP) in 2002 and the M.Sc. and Ph.D degrees in Computer Science from National Institute for Astrophysics, Optics and Electronics (INAOE) in 2004 and 2008, respectively. Since 2017, he is researcher of the Mexican National Council for Science and Technology (CONACYT). Currently, he has been involved in the design and development of cryptographic systems, reconfigurable architectures, software radio platforms, FPGA implementations, and application specific hardware acceleration.

Luis Alberto Morales-Rosales graduated in 2009 from the National Institute of Optical and Electronic Astrophysics as a Ph.D. in Computer Sciences. He is part of the Mexican National System of Researchers, distinguished as Level 1. Currently, he works at Michoacán University of San Nicolás de Hidalgo (UMSNH) as Professor Conacyt. His research includes Computational Models for Transportation Engineering, Intelligent Computing, Security and Distributed Systems.

Received: February 29, 2024; Accepted: July 20, 2024.