EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

TECNOLÓGICO NACIONAL DE MÉXICO

INSTITUTO TECNOLÓGICO DE CULIACÁN

ACCESO REMOTO MEDIANTE IDENTIFICACIÓN FACIAL IMPLEMENTADO EN JETSON NANO Y TENSORFLOW LITE

TESIS

PRESENTADA ANTE EL DEPARTAMENTO ACADÉMICO DE ESTUDIOS DE POSGRADO DEL INSTITUTO TECNOLÓGICO DE CULIACÁN EN CUMPLIMIENTO PARCIAL DE LOS REQUISITOS PARA OBTENER EL GRADO DE

MAESTRO EN CIENCIAS DE LA COMPUTACIÓN

POR:

JOSÉ MISAEL BURRUEL ZAZUETA
INGENIERO EN MECATRÓNICA

DIRECTOR DE TESIS:

MC. GLORIA EKATERINE PERALTA PEÑUÑURI

CULIACÁN, SINALOA 15 de agosto del 2021

"2021, Año de la Independencia"

Culiacán, Sin., 10 de Agosto del 2021

DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN
OFICIO: DEPI/261/8/2021

ASUNTO: **Autorización Impresión**

C. JOSÉ MISAEL BURRUEL ZAZUETA
ESTUDIANTE DE LA MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN
PRESENTE.

Por medio de la presente y en virtud de que ha completado los requisitos para el examen de grado de **Maestro en Ciencias de la Computación**, se concede autorización para la impresión de la tesis titulada: **"ACCESO REMOTO MEDIANTE IDENTIFICACIÓN FACIAL IMPLEMENTADO EN JETSON NANO Y TENSORFLOW LITE"** bajo la dirección del(a) **M.C. Gloria Ekaterine Peralta Peñúñuri.**

Sin otro particular reciba un cordial saludo.

ATENTAMENTE
Excelencia en Educación Tecnológica®

M.C. MARÍA ARACELY MARTÍNEZ AMAYA
JEFE(A) DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA
TECNOLÓGICO NACIONAL DE MÉXICO
INSTITUTO TECNOLÓGICO DE CULIACÁN
DEPARTAMENTO DE DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

C.c.p. archivo

MAMA/lucy *

"ACCESO REMOTO MEDIANTE IDENTIFICACIÓN FACIAL IMPLEMENTADO EN JETSON NANO Y TENSORFLOW LITE"

Tesis presentada por el(a):

C. JOSÉ MISAEL BURRUEL ZAZUETA

Aprobada en contenido y estilo por:

M.C. Gloria Ekaterine Peralta Peñúñuri
Director de Tesis

Dr. Héctor Rodríguez Rangel
Secretario

Dra. María Lucía Barrón Estrada
Vocal -1

Dr. Ramón Zatarain Cabada
Vocal -2

M.C. María Aracely Martínez Amaya
Jefe(a) de la División de Estudios de Posgrado e Investigación

Dedicatoria

A mi **mamá**.

Agradecimientos

Agradezco principalmente a mi **mamá** Guadalupe Zazueta por ser un ejemplo a seguir y no dejarme vencer por las adversidades, por demostrarme con hechos que en la vida hay que esforzarse para lograr cumplir nuestras metas, por enseñarme que los valores son pilares para crecer como persona y por hacer de mí el hombre que soy ahora.

A mi **abuela** Clementina Felix por siempre cuidar de mi como si aun fuera un niño.

A mis **suegros** Mireya López y Humberto Gurrola por aconsejarme y siempre brindarme confianza y mucho cariño.

A mi **hermana** Michelle Montoya por ser mi motivación para superarme día con día y por compartir los momentos más bellos de mi vida.

A mi **novia** Valeria Gurrola por su paciencia, comprensión y amor que me animaban en los malos momentos, por enseñarme el lado más hermoso de la vida y permitirme compartirla contigo.

A mi **asesora** de tesis MC. Gloria Peralta y al Dr. Héctor Rodríguez por guiarme pacientemente para lograr este trabajo que representa dos años de mucho esfuerzo y dedicación.

A mis **profesores** Dra. Lucia Barrón por regañarme siempre con cariño y ser un ejemplo a seguir, Dr. Ramón Zatarain por su comprensión y su buen humor, Dr. Ricardo Quintero por la pasión que demuestra en cada una de sus clases, Dr. Víctor González por enseñarnos temas tan complejos de una forma tan sencilla, al Ing. José Alcaraz y al MC. Dagoberto Tolosa por siempre apoyarme incondicionalmente en el área de Metal-Mecánica y por último a Lucy López por invitarme a formar parte de la MCC y apoyarnos en todos los tramites.

A mis **compañeros** Víctor Bátiz, Eduardo Huerta, Rafael Imperial, Daniel Leyva, Manuel Medrano, Saúl Palazuelos, Marcos Plata, Óscar Sandoval y Rafael Zavala por el apoyo moral, la solidaridad y por brindarme cada uno su granito de arena para lograr esta meta.

Al **Instituto Tecnológico de Culiacán** por mi formación profesional en licenciatura y maestría.

Al **Consejo Nacional de Ciencia y Tecnología** por financiar mis estudios de maestría.

Declaración de Autenticidad

Por la presente declaro que, salvo cuando se haga referencia específica al trabajo de otras personas, el contenido de esta tesis es original y no se ha presentado total o parcialmente para su consideración para cualquier otro título o grado en esta o cualquier otra Universidad. Esta tesis es resultado de mi propio trabajo y no incluye nada que sea resultado de algún trabajo realizado en colaboración, salvo que se indique específicamente en el texto.

José Misael Burruel Zazueta. Culiacán, Sinaloa, México, 2021.

Resumen

Existen diversas tecnologías empleadas para el desbloqueo automático de puertas, algunas de estas técnicas utilizan sistemas biométricos para analizar corporalmente al usuario y de esa forma asegurar que es una persona deseable. Un ejemplo muy común es la lectura de huellas digítales que se utilizan en los bancos, tiendas departamentales, laboratorios o muchos otros edificios que requieren de altos niveles de seguridad para evitar que personas no deseables logren tener acceso.

En la actualidad, la comunidad científica ha hecho grandes contribuciones a la seguridad habitacional incorporando técnicas de inteligencia artificial a los sistemas domóticos que buscan disminuir el esfuerzo humano en tareas complejas o tediosas.

En este trabajo, se propone un método de identificación facial para su implementación en sistemas de cerraduras biométricas. Este método está basado principalmente en la utilización de una red neuronal convolucional desarrollada por Google llamada Facenet. Así mismo, se propone un sistema embebido de cerradura electrónica para implementarse mediante el identificador facial. A diferencia de otros proyectos descritos en el estado del arte, el sistema de identificación facial logra más del 95 % de tasa de reconocimiento correcto y un alto rendimiento.

De igual forma, se propone un sistema embebido conformado por diferentes dispositivos electrónicos, que en su conjunto, generan un prototipo capaz de ejecutar el sistema de reconocimiento facial, que comparado con los trabajos similares del estado del arte, es mejor en diseño y adaptación a las necesidades del usuario según los resultados obtenidos.

El sistema se ha desarrollado para el control de acceso mediante el uso de identificación facial y se realizaron pruebas tomando parámetros de medición propuestos por autores altamente reconocidos por sus aportes en la investigación del aprendizaje profundo. Este trabajo brinda grados de evaluación competitivas, portabilidad y genera líneas de investigación para futuros proyectos.

Palabras Clave

- Aprendizaje profundo
- Cerradura biométrica
- Detección de realidad
- Embedding
- FaceNet
- Identificación facial
- Jetson Nano
- Redes neuronales convolucionales

Índice general

Índice de figuras x

Índice de tablas xii

1. Introducción 1
- 1.1. Definición del problema . . . 2
- 1.2. Hipótesis . . . 2
- 1.3. Objetivo . . . 2
  - 1.3.1. Objetivos específicos . . . 3
- 1.4. Justificación . . . 3
- 1.5. Estructura de la tesis . . . 4

2. Marco teórico 6
- 2.1. Sistemas de cerraduras de puertas . . . 6
  - 2.1.1. Sistema de cerradura mecánica . . . 6
  - 2.1.2. Sistema de cerradura de contraseña . . . 7
  - 2.1.3. Sistema de cerradura RFID . . . 8
  - 2.1.4. Sistema de cerradura biométrica . . . 8
  - 2.1.5. Sistema de cerradura OTP . . . 9
  - 2.1.6. Sistema de cerradura basado en criptografía . . . 9
  - 2.1.7. Sistema de cerradura inalámbrico . . . 10
  - 2.1.8. Sistema de cerradura basado en IoT . . . 10
- 2.2. Inteligencia artificial . . . 11
- 2.3. Aprendizaje máquina . . . 12
- 2.4. Aprendizaje profundo . . . 14
  - 2.4.1. Redes neuronales artificiales . . . 15
  - 2.4.2. Perceptrón . . . 16
  - 2.4.3. Redes neuronales multicapa . . . 17
  - 2.4.4. Redes neuronales convolucionales . . . 17
- 2.5. Reconocimiento facial . . . 19
  - 2.5.1. Visión computacional . . . 19
  - 2.5.2. FaceNet . . . 19
  - 2.5.3. Pérdida de tripletes (Triplet Loss) . . . 21
- 2.6. Tecnologías utilizadas . . . 22
  - 2.6.1. Python . . . 22

2.6.2. TensorFlow . . . 22
2.6.3. Arduino IDE . . . 23
2.6.4. SolidWorks . . . 23
2.7. Hardware utilizado . . . 24
2.7.1. Jetson Nano . . . 24
2.7.2. Arduino Uno . . . 25
2.7.3. Impresora 3D . . . 26
2.8. Críterios de evaluación . . . 26

3. Estado del Arte **28**
3.1. Cerraduras electrónicas . . . 28
3.2. Reconocimiento facial . . . 30
3.2.1. FaceNet . . . 34
3.2.2. Deep Face . . . 35
3.2.3. VGG Face . . . 36

4. Diseño de sistema embebido **38**
4.1. Diseño de hardware . . . 38
4.1.1. Diseño del prototipo . . . 40
4.1.1.1. Componentes electrónicos . . . 41
4.1.1.2. Carcasa para componentes electrónicos . . . 42
4.1.2. Ensamblaje de componentes electrónicos y carcasas . . . 43
4.1.3. Impresión de la carcasa . . . 44
4.2. Desarrollo del Sistema Identificador Facial . . . 45
4.2.1. Análisis de requisitos . . . 46
4.2.1.1. Requisitos no funcionales . . . 46
4.2.1.2. Requisitos funcionales . . . 46
4.2.2. Actores . . . 47
4.2.3. Casos de uso . . . 47
4.2.4. Diagrama de contexto . . . 48
4.2.5. Arquetipos . . . 48
4.2.6. Arquitectura . . . 49
4.2.7. Aplicación . . . 51
4.2.7.1. Creación de la base de datos . . . 51
4.2.7.2. Reconocimiento facial . . . 52

5. Pruebas y resultados **55**
5.1. Ensamblaje de prototipo . . . 55
5.2. Sistema Identificador Facial . . . 56

6. Conclusiones **62**
6.1. Conclusiones . . . 62
6.2. Aportaciones . . . 63
6.3. Trabajo Futuro . . . 64

Bibliografía **65**

Índice de figuras

2.1. Sistema de cerradura mecánica común. 7
2.2. Sistema de cerradura de contraseña. 7
2.3. Sistema de cerradura RFID. 8
2.4. Sistema de cerradura con lector de huellas digitales. 9
2.5. Sistema de cerradura OTP. 9
2.6. Sistema de cerradura inalámbrico. 10
2.7. Sistema de cerradura basado en IoT. 11
2.8. Relación entre el aprendizaje profundo, el aprendizaje máquina y la inteligencia artificial (LeCun et al., 2015). 15
2.9. Arquitectura típica de tres capas de una red neuronal artificial de propagación hacia adelante (Jain et al., 1996). 16
2.10. Representación matemática del perceptrón (Adeli & Yeh, 1989). 16
2.11. Arquitectura clásica de una CNN (LeCun et al., 2015). 18
2.12. Estructura de la CNN FaceNet (Schroff et al., 2015). 19
2.13. Módulo Inception (Szegedy et al., 2015). 20
2.14. Distancias entre la imagen base, la imagen positiva y la imagen negativa (Schroff et al., 2015). 21
2.15. Ventana principal de SolidWorks. 24
2.16. Dispositivo Jetson Nano. 25
2.17. Dispositivo Arduino Uno. 25
2.18. Impresora 3D. 26

4.1. Componentes de hardware utilizados en la cerradura biométrica. 39
4.2. Diagrama de interacción de los dispositivos de hardware en el sistema identificador facial. 40
4.3. Diseños de componentes electrónicos en SolidWorks. 41
4.4. Diseños en SolidWorks de carcasa para Jetson Nano y ventilador. 42
4.5. Diseños en SolidWorks de carcasa para pantalla LCD. 43
4.6. Diseño en SolidWorks de sujetador de módulo de cámara. 43
4.7. Ensamblaje de componentes electrónicos y carcasa. 44
4.8. Proceso de impresión de la base de la Jetson Nano. 45
4.9. Diagrama de casos de uso 47
4.10. Diagrama de contexto 48
4.11. Relación de arquetipos 49
4.12. Arquitectura del sistema 50

4.13. Extracción de la arquitectura para la creación de la base de datos 51
4.14. Resultados del proceso de aumento de datos. 52
4.15. Extracción de la arquitectura para la creación de la base de datos 53
4.16. Ejemplo de una detección facial 53
4.17. Resultados de la detección de realidad en fotografías. 54
4.18. Ejemplo de un reconocimiento facial exitoso. 54

5.1. Ensamblaje completo de prototipo. 55
5.2. Prototipo y puerta para pruebas. 56
5.3. Pruebas realizadas al sistema identificador facial. 57

Índice de tablas

3.1. Lista de componentes utilizados por los autores. 29
3.2. Resultados obtenidos de los intentos de lectura de huella digítal. 29
3.3. Lista de componentes utilizados por los autores Falohun et al., 2012. 30
3.4. Resultados obtenidos por Bhattacharyya et al., 2009. 31
3.5. Criterios de evaluación y niveles de clasificación de los sistemas biométricos. 33
3.6. Resultados obtenidos por los autores. 34

5.1. Relación de identificación real y predicha por el sistema a usuarios presenciales. 57
5.2. Relación de identificación real y predicha por el sistema a usuarios virtuales. . 58
5.3. Matriz de confusión. 59
5.4. Resultados obtenidos de las mediciones realizadas. 60
5.5. Comparación de componentes utilizados por otros autores. 61
5.6. Comparación de otros sistemas con el sistema propuesto en este trabajo. . . . 61

Capítulo 1

Introducción

El robo a casa habitación ha estado en la quinta posición de incidencia delictiva en México por muchos años, incluso por encima de robo total de automóvil, según cifras del INEGI (INEGI, 2021). Poca vigilancia, sistemas de seguridad deficientes y otros factores han contribuido a que los robos en propiedades privadas tengan tasas de incidencia alarmantes. Gracias a la domótica es posible generar más y mejores sistemas de seguridad que brinden protección a nuestros hogares.

La domótica es un conjunto de diferentes tecnologías aplicadas a la monitorización, control y automatización de sistemas y dispositivos en el hogar. Los principales objetivos de la domótica son mejorar la seguridad personal y patrimonial del hogar, incrementar el confort y tener una gestión eficiente del uso de la energía (García & Vega, 2018). Algunas de estas tecnologías se aplican en los sistemas de cerradura de puertas biométricas mediante diferentes técnicas, e.g. huellas digitales, iris ocular, reconocimiento facial, entre otros.

El reconocimiento facial se ha utilizado en diversos campos de la computación con distintos fines de aplicación, tales como en las redes sociales para identificar a las personas que aparecen en fotografías, en aplicaciones móviles de teléfonos inteligentes, la educación para identificar emociones mediante expresiones faciales y sobre todo en herramientas de seguridad. El principal objetivo de este proyecto es utilizar el reconocimiento facial como herramienta esencial para diseñar un sistema de seguridad en conjunto con un dispositivo Jetson Nano. Se pretende cumplir este objetivo utilizando redes neuronales convolucionales para el tratamiento y procesamiento de las imágenes.

1.1. Definición del problema

La utilización de contraseñas, tarjetas con chip, llaves o un sinfín de objetos reemplazables son poco prácticos y carecen de la seguridad necesaria para garantizar que una persona no deseada logre el acceso a un lugar específico. Para monitorizar de una forma operativa el acceso de personas a un área restringida y tener certeza de que se trata de un individuo deseable, es necesario instalar cerraduras biométricas en las puertas de acceso a dicha área.

Con la ayuda de una cerradura biométrica, específicamente, de una cerradura con reconocimiento facial, se garantizaría que los accesos serian de personas deseables sin hacer contacto físico con el usuario, ya que, el uso de otros sistemas biométricos (e.g. lectura de huella dactilar o lectura de iris ocular) requieren de una cooperación mayor por parte del usuario o incluso algunas llegan a ser invasivas.

El reconocimiento facial es una herramienta que mejora la tarea de identificación de personas. Pese a que este tipo de tecnologías ya se encuentran aplicadas en varios campos, tienen deficiencias en aspectos de altos costos, efectividad, velocidad y/o portabilidad, para alcanzar mejores estándares de calidad se requiere del uso de nuevas tecnologías que utilizan el reconocimiento facial.

1.2. Hipótesis

Un sistema de identificación facial desarrollado con técnicas de inteligencia artificial aporta una herramienta de alta calidad para ser implementado en sistemas embebidos de cerradura de puertas.

1.3. Objetivo

Diseñar e implementar un control de acceso remoto mediante identificación facial utilizando inteligencia y visión artificial en un dispositivo Jetson Nano.

1.3.1. Objetivos específicos

- Desarrollar un sistema que identifique rostros en video en tiempo real.
- Identificar que los rostros mostrados en video sean de personas reales.
- Validar que los rostros detectados en el sistema tienen o no acceso.
- Implementar un sistema completo en un prototipo de acceso remoto basado en una computadora Jetson Nano.

1.4. Justificación

La restricción de accesos es un tema de seguridad básica que se lleva a cabo desde los inicios de la humanidad. Con la intención de cubrir estas necesidades, se han desarrollado mecanismos de seguridad para garantizar que entes no deseados tengan acceso a diferentes espacios, algunos de estos son: cerraduras, cortinas metálicas, candados y barricadas. La mayor parte de estos aparatos que actualmente se siguen utilizando en la vida cotidiana, funcionan con llaves construidas con patrones que, presuponen singularidad para que esa cerradura o candado no pueda ser abierto con una llave que no corresponda a su patrón interno.

Con el avance de la tecnología y en específico, de la inteligencia artificial (IA), se han desarrollado proyectos que combinan los sistemas de seguridad con técnicas de inteligencia artificial, esto con el fin de mejorar los sistemas rudimentarios y al mismo tiempo facilitar su uso. Este avance por supuesto se ha dado de manera sustancial, por ejemplo, sustituyendo las llaves comunes, por tarjetas electrónicas o contraseñas alfanuméricas, al mismo tiempo que se sustituyeron los candados y cerraduras, por dispositivos sensoriales que identifican las tarjetas y contraseñas antes mencionadas.

Si bien fueron muchos los logros en el ámbito de seguridad con la ayuda de la IA, existen algunos defectos que no garantizan el no acceso de personas indeseadas. Extraviar u olvidar tarjetas electrónicas o hackeo de contraseñas, son dos ejemplos claros de aspectos que no se logran cubrir con los dispositivos descritos anteriormente. Con la ayuda del aprendizaje profundo y el tratamiento masivo de imágenes, se puede lograr cubrir estos defectos con el uso

del reconocimiento facial de las personas que quieren abrir una puerta. El aprendizaje profundo garantiza en más del 95 % que la persona sometida a la identificación facial es reconocida satisfactoriamente, por lo que el desarrollo de estas técnicas es de suma importancia para la comunidad tecnológica en las áreas de visión artificial, aprendizaje profundo y seguridad.

Aunado a todo lo anteriormente descrito, no es suficiente el desarrollo de las técnicas de aprendizaje profundo y visión artificial, si no pueden ser ejecutadas en dispositivos con altos estándares portabilidad y que logren mantener los niveles de seguridad necesarios. La Jetson Nano es un dispositivo computacional que provee estas características y además también es altamente potente comparado con el resto de los dispositivos en el mercado. Mezclando la Jetson Nano con un sistema de reconocimiento facial se obtendrá un dispositivo capaz de bloquear o desbloquear una puerta mediante la identificación del rostro del usuario.

1.5. Estructura de la tesis

El presente trabajo esta divido en 6 capítulos incluyendo este capítulo de Introducción y se organiza de la siguiente manera:

- Capítulo 2 - Marco Teórico: Se describen los conceptos elementales para el desarrollo del proyecto como lo son la inteligencia artificial, aprendizaje máquina, aprendizaje profundo, visión artificial, reconocimiento facial y las tecnologías utilizadas las cuales fueron las herramientas primarias para la generación e implementación del proyecto.
- Capítulo 3 - Estado del Arte: Se muestran las tecnologías que se han desarrollado hasta la actualidad en torno y en relación con la investigación desarrollada en el presente escrito.
- Capítulo 4 - Diseño de sistema embebido: Se explican las decisiones de desarrollo del sistema desde la arquitectura, la codificación, la obtención de resultados y el diseño del prototipo.
- Capítulo 5 - Pruebas y resultados: Se presentan los resultados obtenidos en la ejecución del sistema de identificación facial relacionados con el ensamblaje del prototipo y el rendimiento del reconocimiento facial.

- Capítulo 6 - Conclusiones: Se realiza un resumen de los métodos utilizados para generar un enfoque a los resultados obtenidos y de esa forma determinar conclusiones acerca del proyecto. Con base en lo concluido se determinan posibles trabajos a futuro con el fin de mejorar o ampliar el panorama del trabajo.

Capítulo 2

Marco teórico

La base conceptual de este trabajo es presentada en este capítulo con el objetivo de definir los conceptos y herramientas utilizados para el desarrollo del proyecto. Este capítulo se divide en cuatro secciones, en la primera sección explican los tipos de cerraduras que existen actualmente y sus diferentes tipos de operación, en la segunda sección se definen los conceptos más importantes que tienen lugar a lo largo del desarrollo del proyecto, partiendo desde lo general a lo particular, i.e., comenzando desde la inteligencia artificial y llegando a la profundidad de las descripciones de perceptrón y redes neuronales convolucionales. En la tercera sección se describen las tecnologías de programación y diseño utilizadas para desarrollar la totalidad del proyecto. En la última sección se describe el hardware empleado para implementar los sistemas y construir el prototipo.

2.1. Sistemas de cerraduras de puertas

Las cerraduras de puertas son sistemas que contribuyen al resguardo de la seguridad de cualquier edificación, estos sistemas pretenden evitar el ingreso de personas no autorizadas a zonas específicas. En Divya & Mathew, 2017 se describen algunas tecnologías empleadas en la seguridad de las cerraduras de puertas.

2.1.1. Sistema de cerradura mecánica

Este tipo de sistemas se componen de dos partes, la cerradura y la llave. Son los más utilizados en todo el mundo por su precio, facilidad de uso e instalación, pero al mismo tiempo

son los sistemas que resultan más débiles a la hora de brindar protección contra ladrones. En la Figura 2.1 se muestra un ejemplo típico de una cerradura mecánica, este tipo de cerraduras es el sistema más comúnmente utilizado en nuestros hogares.

Figura 2.1: Sistema de cerradura mecánica común.

2.1.2. Sistema de cerradura de contraseña

Este sistema es una combinación de hardware y software en su máxima expresión. En este diseño, un dispositivo se bloquea mediante un código (contraseña). El código se puede configurar según el deseo del usuario. El código electrónico bloquea el sistema al cambiar al modo de alarma cuando se ingresa un código incorrecto y permite al usuario iniciarlo solo cuando se ingresa un código en la secuencia correcta (Oke Alice et al., 2013). En la Figura 2.2 se muestra un ejemplo de una cerradura accionada por contraseña numérica.

Figura 2.2: Sistema de cerradura de contraseña.

2.1.3. Sistema de cerradura RFID

La identificación por radiofrecuencia (RFID, por sus siglas en inglés) es una tecnología fundamental y económica que permite la transmisión inalámbrica de datos. Además, se implementa un sistema de bloqueo de puerta digital y se rige por un lector RFID que autentica y válida al usuario y abre la puerta automáticamente. También mantiene el registro de entrada y salida del usuario (Verma & Tripathi, 2010). En la Figura 2.3 se muestra un ejemplo de una cerradura accionada por una tarjeta RFID.

Figura 2.3: Sistema de cerradura RFID.

2.1.4. Sistema de cerradura biométrica

El sistema biométrico otorga acceso a los usuarios autorizados mediante la verificación de sus características físicas o de comportamiento únicas, como huellas dactilares, reconocimiento facial, reconocimiento de voz, detector de venas, escáner de iris, etc. Estos sistemas de bloqueo funcionan escaneando los datos biométricos y luego convirtiéndolos en una plantilla numérica que se guardará por primera vez. Luego, la próxima vez que alguien intente acceder a la puerta utilizando sus datos biométricos, se comparará con el valor guardado previamente (Divya & Mathew, 2017). En la Figura 2.4 se muestra una cerradura biométrica accionada por un lector de huellas digitales.

Figura 2.4: Sistema de cerradura con lector de huellas digitales.

2.1.5. Sistema de cerradura OTP

La contraseña de un solo uso (OTP, por sus siglas en inglés) es conocida como contraseña desechable que será válida solo una vez y se producirá una cada vez que intente acceder al sistema. Esta contraseña consiste en una cadena de caracteres numéricos o alfanuméricos que se producen de forma automática. El servidor generará OTP y se enviará al sistema, así como al teléfono registrado. El sistema se desbloquea si la OTP ingresada por el usuario coincide con la OTP recibida en el sistema (Divya & Mathew, 2017). En la Figura 2.5 se muestra una cerradura que funciona con una contraseña de un solo uso.

Figura 2.5: Sistema de cerradura OTP.

2.1.6. Sistema de cerradura basado en criptografía

El cifrado es un método para evitar el acceso no autorizado donde la información original se convierte a otra forma. El software de criptografía se utiliza al ingresar la contraseña para evitar la piratería por parte de un empleador corrupto. Aquí se establece una contraseña dentro del chip del microcontrolador. Ahora, esta contraseña establecida se cifrará con otros datos

que se le darán al usuario como contraseña para desbloquear el sistema (Divya & Mathew, 2017).

2.1.7. Sistema de cerradura inalámbrico

El sistema consta de una cerradura de puerta y un dispositivo informático móvil. La cerradura de la puerta consta de un dispositivo de bloqueo, un dispositivo de comunicación de campo cercano (NFC, por sus siglas en inglés) y un microcontrolador, el dispositivo informático móvil consta de un dispositivo NFC, una pantalla y una aplicación móvil. La aplicación móvil genera un código que se transmite a través de la señal NFC en respuesta a la comunicación con el servidor. El microcontrolador después de recibir este código desconecta el dispositivo de bloqueo determinando que el código recibido contiene los datos correctos para desbloquear la puerta (Divya & Mathew, 2017). En la Figura 2.6 se muestra un sistema de cerradura accionado por un dispositivo inalámbrico.

Figura 2.6: Sistema de cerradura inalámbrico.

2.1.8. Sistema de cerradura basado en IoT

El Internet de las cosas (IoT, por sus siglas en inglés) se puede definir como la conexión de varios tipos de objetos como teléfonos inteligentes, computadoras personales y tabletas a Internet, lo que brinda un tipo de comunicación muy novedosa entre las cosas y las personas. Con la introducción de IoT, la investigación y el desarrollo de la automatización del hogar se están volviendo populares en los últimos días. Muchos de los dispositivos están controlados y monitorizados para ayudar al ser humano (Pavithra & Balakrishnan, 2015). En la Figura

2.7 se observa un sistema de cerradura basado en internet de las cosas, siendo accionado por un teléfono inteligente.

Figura 2.7: Sistema de cerradura basado en IoT.

2.2. Inteligencia artificial

Para definir un concepto tan complejo como lo es la inteligencia artificial, es necesario dar un recorrido en el pasado para comprender las ideas que dieron forma a este tema tan extenso y con tanto auge en la actualidad.

A fines de la década de 1940, un joven autor llamado Isaac Asimov comenzó a escribir una serie de historias y novelas sobre robots. A lo largo de su vida, Asimov creyó que sus Tres Leyes eran más que un simple recurso literario; sentía que los científicos e ingenieros involucrados en robótica e investigadores de inteligencia artificial (IA) se habían tomado sus leyes en serio (McCauley, 2007).

¿Cuáles son las tres leyes? Más allá de escribir historias sobre robots "buenos", Asimov los imbuyó primero con tres leyes explícitas expresado en forma impresa en la historia, "Runaround" (Asimov, 2004):

- 1. "Un robot no puede dañar a un ser humano o, a través de inacción, permitir que un ser humano sufra daños".
- 2. "Un robot debe obedecer las órdenes que le den los seres humanos. Excepto donde tales órdenes entrarían en conflicto con la Primera Ley".
- 3. "Un robot debe proteger su propia existencia, siempre que tal protección no entra en conflicto con la Primera o Segunda Ley".

La primera vez que el término inteligencia fue asociado a una máquina fue en los años

50 por Alan Turing, padre de la computación. Turing afirmaba que si una máquina tenía la capacidad de pensar entonces era inteligente (Turing, 1950).

La palabra Inteligencia Artificial fue acuñada oficialmente unos seis años después, cuando en 1956 Marvin Minsky y John McCarthy organizaron el Proyecto de Investigación de Verano de Dartmouth sobre Inteligencia Artificial (DSRPAI) de aproximadamente ocho semanas de duración en la Universidad de Dartmouth en New Hampshire. Este taller, que marca el comienzo de AI Spring y fue financiado por la Fundación Rockefeller, reunió a quienes más tarde serían considerados como los padres fundadores de IA (Buchanan, 2005).

De la decada de los 60 a los 90 el objetivo del estudio de la Inteligencia Artificial ya no era crear un robot tan inteligente como un humano, sino utilizar algoritmos, heurísticas y metodologías basadas en las formas en que el cerebro humano resuelve problemas (Coppin, 2004).

A pesar de que han pasado 70 años de este primer acercamiento a la inteligencia artificial aún no existe una definición establecida del término en su conjunto. La siguiente definición interpreta a la inteligencia artificial como campo de estudio y no como a un concepto. “Es la ciencia y la ingeniería para fabricar máquinas inteligentes, especialmente programas informáticos inteligentes. Está relacionado con la tarea similar de usar computadoras para comprender la inteligencia humana, pero la IA no tiene que limitarse a métodos que son biológicamente observables” (McCarthy, 2007).

La IA actualmente logra comprender una gran diversidad de subcampos, que van desde lo general (aprendizaje y percepción) hasta lo específico, como jugar al ajedrez, demostrar teoremas matemáticos, escribir poesía, conducir un automóvil en una calle concurrida y diagnosticar enfermedades. La IA es sobresaliente para cualquier tarea especulativa; es verdaderamente un campo universal (Nilsson, 1996).

2.3. Aprendizaje máquina

Según Mohri et al., 2018 el aprendizaje máquina se puede definir ampliamente como procedimientos computacionales que utilizan la experiencia para aumentar el rendimiento o hacer predicciones más precisas. Aquí, la experiencia hace referencia a la información anti-

gua disponible para el usuario, que generalmente toma la forma de información electrónica almacenada y preparada para su análisis.

Básicamente, el aprendizaje máquina utiliza algoritmos para extraer información de datos sin procesar y representarla en algún tipo de modelo. Se usa este modelo para inferir cosas sobre otros datos que aún no se han modelado (Patterson & Gibson, 2017).

Actualmente el aprendizaje máquina se puede encontrar en muchos campos de aplicación y que con el paso del tiempo, cada vez son más las tareas que se pueden resolver con su uso. A continuación se presentan algunas de las aplicaciones más comunes del aprendizaje máquina (X.-D. Zhang, 2020):

- Reconocimiento de voz.
- Diagnósticos médicos.
- Juegos, e.g. ajedrez.
- Control de vehículos no asistidos.
- Clasificación de textos.
- Visión artificial, e.g. reconocimiento facial, reconocimiento de objetos.

En Mohri et al., 2018 se asignan cada una de las tareas a un tipo de problema que el aprendizaje máquina es capaz de resolver, en esta misma referencia se definen los siguientes conceptos de cada problema:

- Clasificación: asignar una categoría a cada artículo. Por ejemplo, la clasificación de documentos permite asignar elementos a categorías como política, negocios, deportes y clima, y la clasificación de imágenes permite asignar elementos a categorías como paisaje, retratos y animales.
- Regresión: predice un valor real para cada elemento. Los ejemplos de regresión integran la predicción de valores de existencias o variaciones de cambiantes económicas. En este problema, la penalización por una predicción errónea es dependiente del tamaño de la diferencia entre los valores verdaderos y predichos, en contraste con el

problema de categorización, donde típicamente no hay idea de cercanía entre numerosas categorías .

- Categorización: ordenar artículos de acuerdo con algún criterio. La búsqueda web es un ejemplo de una taxonomía legítima, como devolver una página web relacionada con la consulta de búsqueda. Muchos otros problemas de clasificación similares ocurren en el contexto de la recuperación de información o el diseño del sistema de procesamiento del lenguaje natural.

- Agrupación: partición de elementos en regiones homogéneas. La agrupación en clústeres se utiliza comúnmente para analizar conjuntos de datos muy grandes.. Por ejemplo, en el tema del análisis de redes sociales, los algoritmos de agrupación tratan identificar la comunidad dentro un gran grupo de personas.

- Reducción de dimensionalidad o aprendizaje múltiple: Convierte la representación original de un elemento en una representación dimensional inferior de esos elementos, conservando algunas propiedades de la representación original. Los ejemplos comunes incluyen el preprocesamiento de imágenes digitales en tareas de visión por computadora.

Si bien, más que problemas se entendería que los conceptos anteriormente descritos pueden definirse como los objetivos del aprendizaje máquina, objetivos que son perseguidos con la finalidad de lograr mejores predicciones en las tareas que se ven involucradas.

2.4. Aprendizaje profundo

El aprendizaje profundo faculta que los modelos computacionales que se constituyen de numerosas capas de procesamiento aprendan representaciones de datos con muchos niveles de abstracción. Estos métodos han optimizado drásticamente el estado de la técnica en reconocimiento de voz, reconocimiento de objetos visuales, detección de objetos y algunos otros dominios, como el descubrimiento de fármacos y la genómica (LeCun et al., 2015).

Actualmente el aprendizaje profundo ha sido una herramienta de alta demanda debido a su capacidad de resolución de diferentes problemas propuestos al aprendizaje máquina,

la realidad es que el aprendizaje profundo abarca una amplia gama de modelos capaces de aprender a realizar diversas tareas propuestas por los diseñadores de dichos modelos. En la Figura 2.8 se observa la relación entre el aprendizaje profundo, el aprendizaje máquina y la inteligencia artificial.

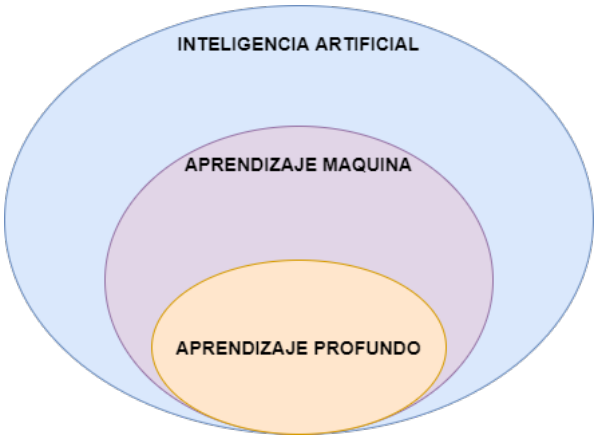


Figura 2.8: Relación entre el aprendizaje profundo, el aprendizaje máquina y la inteligencia artificial (LeCun et al., 2015).

2.4.1. Redes neuronales artificiales

Las redes neuronales artificiales se inspiran en la arquitectura de neuronas biológicas como el cerebro humano. El cerebro humano está formado por una gran cantidad de neuronas interconectadas. Cada neurona es una célula que realiza tareas simples como responder a las señales entrantes. Al igual que las redes neuronales biológicas, las redes neuronales artificiales son una interconexión de nodos. Cada red neuronal tiene tres componentes importantes: las características de los nodos, la topología de la red y las reglas de aprendizaje (Zou et al., 2008). En la Figura 2.9 se muestra un ejemplo típico de una red neuronal artificial de propagación hacia adelante.

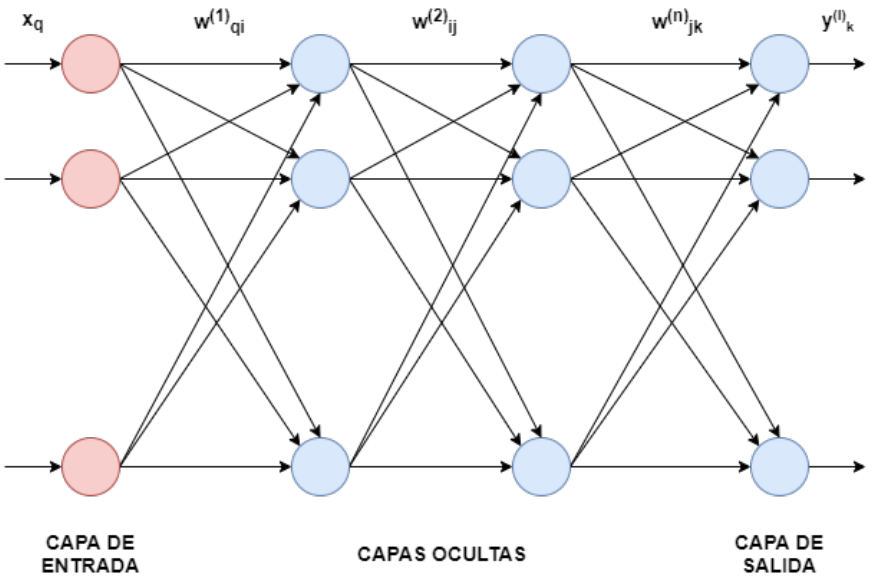


Figura 2.9: Arquitectura típica de tres capas de una red neuronal artificial de propagación hacia adelante (Jain et al., 1996).

2.4.2. Perceptrón

El perceptrón es un dispositivo que puede responder "sí" o "no" a un problema en el dominio que se está considerando (Figura 2.10). Se modela un perceptrón P como una entidad de cuatro tuplas (S, W, F, T) definida de la siguiente manera (Adeli & Yeh, 1989):

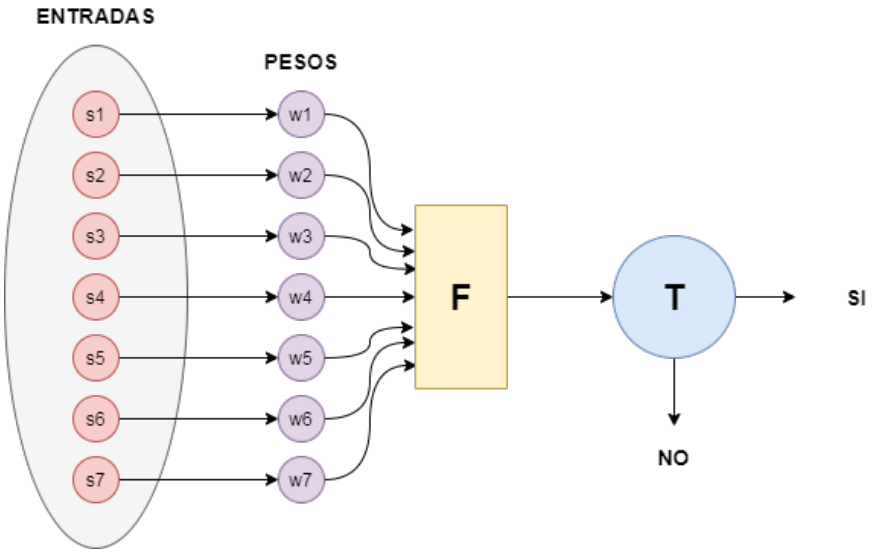


Figura 2.10: Representación matemática del perceptrón (Adeli & Yeh, 1989).

- S significa sensor. Los sensores se postulan como un conjunto de predicados, cada uno de los cuales proporciona una observación y genera datos sobre un subdominio del dominio de observación.

- W significa peso. Los datos obtenidos de los sensores se multiplican por sus pesos correspondientes para producir la interpretación de cada fenómeno observado.

- F es una función que recopila todos los datos ponderados para producir una medición adecuada del impacto de los fenómenos observados.

- T es un umbral con nombre constante. Al comparar T con el valor de salida de la función F, el perceptrón responde "sí" si f produce un valor mayor que T, de lo contrario el perceptrón responde "no".

2.4.3. Redes neuronales multicapa

Una red neuronal multicapa es una red en capas que consta de una capa de entrada, una capa de salida y al menos una capa de elementos de procesamiento no lineales. Los elementos de procesamiento no lineal, que suman las señales entrantes y generan señales de salida de acuerdo con algunas funciones predefinidas, se denominan neuronas o nodos (Srinivasan et al., 1994).

En la Figura 2.9 se muestra un ejemplo de una red neuronal multicapa, la cual consta de una capa oculta o de elementos de procesamientos no lineales.

2.4.4. Redes neuronales convolucionales

Las redes neuronales convolucionales (CNN, por sus siglas en inglés)(Figura 2.11) están diseñadas para transformar datos que vienen en forma de varias matrices, por ejemplo, una fotografía a color compuesta por tres matrices de dos dimensiones que tienen en su interior intensidades de píxeles en los tres canales de color. Muchas modalidades de datos se encuentran en forma de matrices múltiples: una dimensión para señales y secuencias, incluido el lenguaje; dos dimensiones para imágenes o espectrogramas de audio; y tres dimensiones para vídeo o imágenes volumétricas. Hay cuatro ideas clave detrás de CNN que aprovechan las

propiedades de las señales naturales: conexiones locales, pesos compartidos, agrupación y el uso de muchas capas (LeCun et al., 2015).

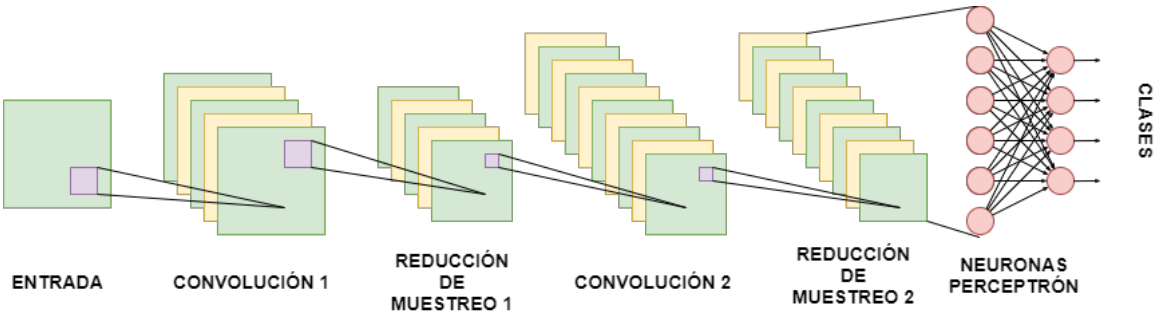


Figura 2.11: Arquitectura clásica de una CNN (LeCun et al., 2015).

Las diferentes capas de una red neuronal convolucional son (Sharma et al., 2018):

- Capa de entrada: la primera capa de cada CNN utilizada es la "capa de entrada", que toma imágenes, cambia su tamaño para pasar a otras capas para la extracción de características.
- Capa de convolución: las siguientes capas son "capas de convolución" que actúan como filtros para las imágenes, por lo tanto, descubren las características de las imágenes y también se utilizan para calcular los puntos de las características de coincidencia durante las pruebas.
- Capa de agrupación: los conjuntos de características extraídos se pasan a la "capa de agrupación". Esta capa toma imágenes grandes y las encoge mientras conserva la información más importante en ellas. Mantiene el valor máximo de cada ventana, conserva los mejores ajustes de cada característica dentro de la ventana.
- Capa de unidad lineal rectificada: la siguiente "Unidad lineal rectificada" o capa ReLU intercambia cada número negativo de la capa de agrupación con 0. Esto ayuda a la CNN a mantenerse matemáticamente estable al evitar que los valores aprendidos se atasquen cerca de 0 o se disparen hacia el infinito.
- Capa completamente conectada: la capa final es la capa completamente conectada que toma las imágenes filtradas de alto nivel y las traduce en categorías con etiquetas.

2.5. Reconocimiento facial

El reconocimiento facial es el proceso de reconocer el rostro de una persona relevante mediante un sistema de visión. Ha sido una herramienta crucial de interacción humano-computadora debido a su uso en sistemas de seguridad, control de acceso, videovigilancia, áreas comerciales e incluso se usa en redes sociales como Facebook (Coşkun et al., 2017).

2.5.1. Visión computacional

El propósito de la visión por computadora es hacer que las computadoras sean capaces de comprender entornos a partir de información visual. Implica una variedad de procesamiento de información inteligente: tanto el procesamiento de patrones para la extracción de símbolos significativos de la información visual como el procesamiento de símbolos para determinar qué representan los símbolos (Shirai, 2012).

2.5.2. FaceNet

FaceNet es una red neuronal convolucional profunda desarrollada por investigadores de Google e introducida alrededor de 2015 para resolver eficazmente los obstáculos en la detección y verificación de rostros (Jose et al., 2019).

La red consta de una capa de entrada por lotes y una CNN profunda seguida de una normalización L2, que da como resultado la incrustación (Embedding) de la cara. A esto le sigue la pérdida de tripletes (Triplet Loss) durante el entrenamiento (Figura 2.12) (Schroff et al., 2015).

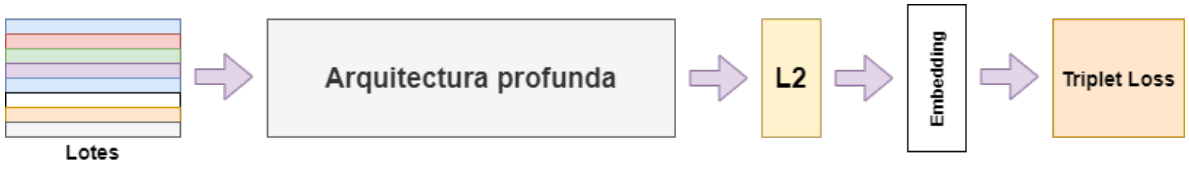


Figura 2.12: Estructura de la CNN FaceNet (Schroff et al., 2015).

La red neuronal profunda del sistema FaceNet tradicional utiliza GoogleNet que también se conoce como red de Inception porque GoogLeNet utiliza múltiples módulos de Inception

(Xu et al., 2020)(Figura 2.13).

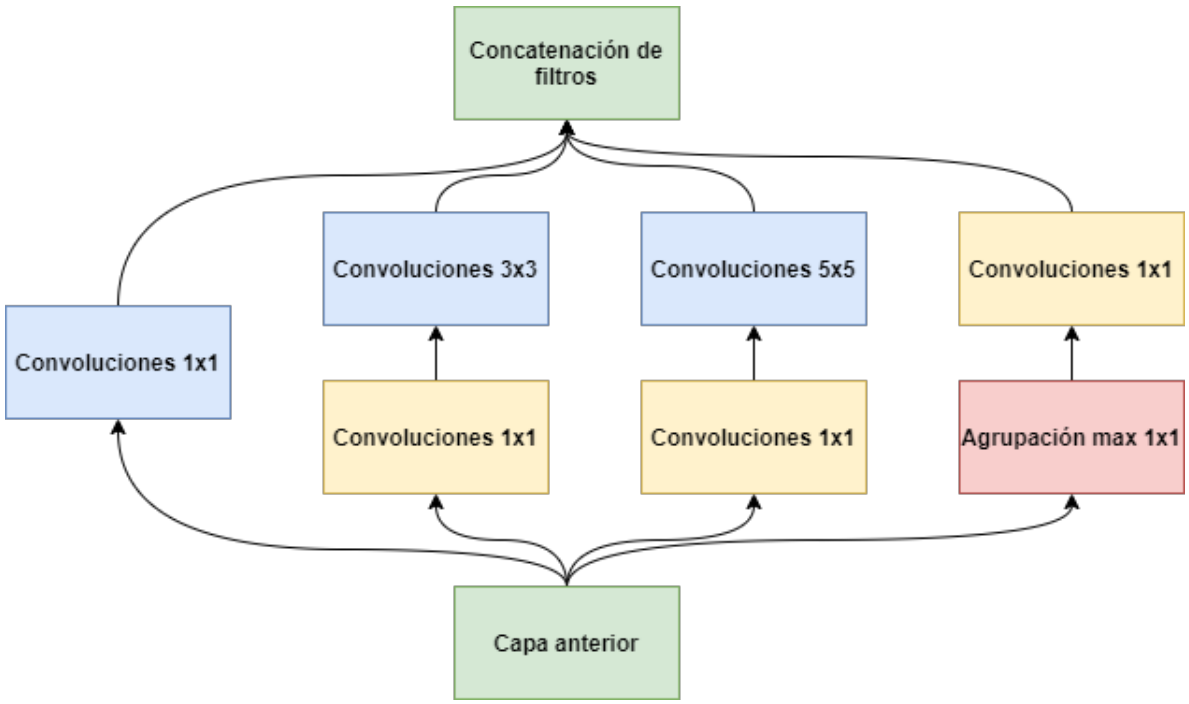


Figura 2.13: Módulo Inception (Szegedy et al., 2015).

La diferencia más notable entre una CNN clásica (Figura 2.11) y la FaceNet, es que en el caso de esta última, la capa de clasificación se elimina para solamente obtener un vector de 128 elementos y no una clasificación de la imagen. Este vector llamado "Embedding", contiene la información esencial de cada rostro y se coloca en un espacio euclidiano. Después de calcular los embeddings de todos los rostros de una base de datos, solamente haría falta comparar el embedding de un rostro nuevo con cada uno de los que se encuentran disponibles en el espacio euclidiano y si la distancia se encuentra por debajo de un umbral con alguno de ellos, se otorgaría la identidad del embedding almacenado al nuevo calculado. Este método descrito anteriormente es el más práctico con el que se puede realizar un reconocimiento facial con la ayuda de la FaceNet.

En todos los experimentos realizados por los autores, se entrenó la CNN usando el Descenso de Gradiente Estocástico (SGD, por sus siglas en ingles) con backprop estándar y AdaGrad. En la mayoría de los experimentos, se comenzó con una tasa de aprendizaje de 0.05 que fue disminuido para finalizar el modelo. Los modelos se inicializan de forma aleatoria, y se entrenan en un clúster de CPU de 1000 a 2000 horas (Schroff et al., 2015).

FaceNet es una CNN pre-entrenada que extrae las características principales de los rostros y las transforma en un vector de 128 dimensiones conocido como “Embedding”. Las CNN convencionales no logran resolver un problema llamado ”One-Shot-Learning”, el cual se le atribuye a la incapacidad de hacer un entrenamiento óptimo con una sola imagen por usuario en la base de datos. En cambio, la FaceNet al ser una CNN pre-entrenada con más de 500 millones de imágenes logra una efectividad del 99.63 % (Schroff et al., 2015), siendo una herramienta muy útil y fácil de usar.

2.5.3. Pérdida de tripletes (Triplet Loss)

Triplet Loss (Figura 2.14) minimiza la distancia entre una imagen base y una imagen positiva, las cuales tienen la misma identidad, y maximiza la distancia entre la imagen base y una imagen negativa de una identidad diferente (Schroff et al., 2015).

Figura 2.14: Distancias entre la imagen base, la imagen positiva y la imagen negativa (Schroff et al., 2015).

FaceNet entrena su salida directamente en una incrustación concisa de 128 dimensiones mediante la aplicación de pérdida basada en tripletes. El Triplet Loss está formado por dos miniaturas de caras iguales y miniaturas que no coinciden y la pérdida tiene como objetivo distinguir entre pares positivos y negativos utilizando un rango de límite (William et al., 2019).

2.6. Tecnologías utilizadas

En esta sección se describen los lenguajes, programas y bibliotecas utilizadas durante el desarrollo del proyecto.

2.6.1. Python

Python es un lenguaje de programación orientado a objetos interpretado e interactivo. Proporciona estructuras de datos de alto nivel como listas y matrices asociativas (llamadas diccionarios), módulos, clases, excepciones y gestión de memoria automática tipados y vinculados dinámicamente. Tiene una sintaxis muy simple y elegante, pero es un poderoso lenguaje de programación de propósito general. Diseñado por Guido van Rossum en 1990. Como muchos otros lenguajes de programación, es gratuito para uso comercial y puede ejecutarse en casi todas las computadoras modernas. El intérprete compila automáticamente el programa Python en un código de bytes independiente de la plataforma y lo interpreta (Sanner et al., 1999).

2.6.2. TensorFlow

TensorFlow es un sistema de aprendizaje automático que opera en entornos heterogéneos y a gran escala. TensorFlow usa diagramas de flujo de datos para representar cálculos, estados compartidos y operaciones que cambian ese estado. Asigna nodos de un gráfico de flujo de datos en varias máquinas de un clúster y entre varios dispositivos de cómputo en una máquina, como procesadores de múltiples núcleos, GPU de uso general y ASIC personalizados conocidos como unidades de procesamiento de tensor (TPU, por sus siglas en inglés). Esta arquitectura proporciona flexibilidad a los desarrolladores de aplicaciones. En proyectos anteriores de “Parameter Server”, la administración de estado compartida se integró en el sistema, pero TensorFlow permite a los desarrolladores experimentar con nuevos algoritmos de capacitación y optimización. TensorFlow admite una amplia variedad de aplicaciones, centrándose en el entrenamiento y la inferencia en redes neuronales profundas (Abadi et al., 2016).

TensorFlow es una combinación de un conjunto de algoritmos y modelos de aprendizaje automático y aprendizaje profundo (o red neuronal). TensorFlow fue fundado y desarrollado por investigadores del equipo de Google Brain y arquitectos de la Organización de Investigación de Inteligencia Artificial de Google con el propósito de investigar el aprendizaje automático y las redes neuronales profundas. (William et al., 2019).

2.6.3. Arduino IDE

Arduino IDE es un software de código abierto que se utiliza principalmente para escribir y compilar código en módulos Arduino. Este es el software oficial de Arduino, e incluso la persona promedio sin conocimientos técnicos previos no puede participar en el proceso de aprendizaje porque el código es demasiado fácil de compilar. Disponible para su uso en sistemas operativos como MAC, Windows y Linux, se ejecuta en la plataforma Java con comandos y funciones incorporados que juegan un papel importante en la depuración, edición y compilación de código en su entorno. Hay muchos módulos Arduino disponibles, incluidos Arduino Uno, Arduino Mega, Arduino Leonardo y Arduino Micro. Cada uno de ellos contiene un microcontrolador integrado que está realmente programado y acepta información en forma de código. El código principal (también conocido como bocetos) generado por la plataforma IDE eventualmente generará un archivo hexadecimal que se pasará al controlador de la placa para su carga. El entorno IDE contiene dos partes principales: editor y compilador. La primera parte se usa para escribir el código requerido y luego compilar el código y cargarlo en la mayoría de los módulos Arduino. Este entorno es compatible con los lenguajes C y C++ (Fezari & Al Dahoud, 2018).

2.6.4. SolidWorks

SolidWorks (Figura 2.15) es un software de modelado de sólidos paramétrico tridimensional desarrollado sobre la base de Windows. Cualquier lenguaje de programación compatible con vinculación e incrustación de objetos y modelo de objetos componentes se puede utilizar como herramientas de desarrollo de SolidWorks, como Visual Basic, Visual Basic Application, VC + +, C#, etc. Visual Basic Application es la herramienta sencilla para gestionar el

desarrollo secundario durante la grabación y edición de macros (Reddy et al., 2016).

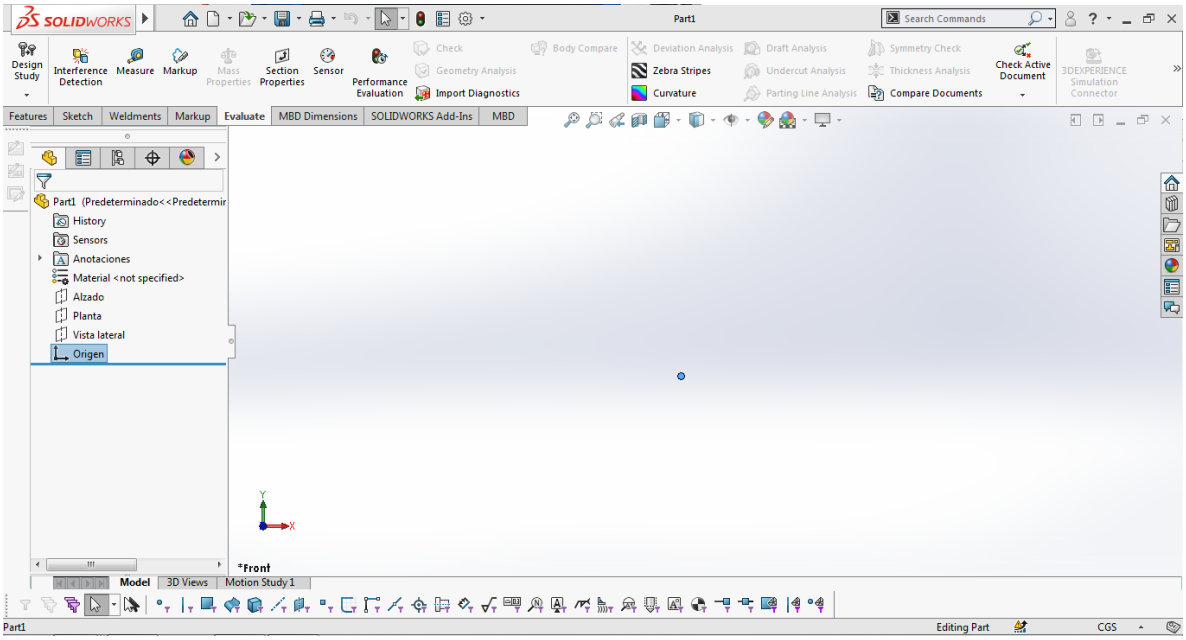


Figura 2.15: Ventana principal de SolidWorks.

2.7. Hardware utilizado

El objetivo de esta sección es mostrar los dispositivos hardware que fueron necesarios para la implementación del sistema embebido de este proyecto.

2.7.1. Jetson Nano

NVIDIA® Jetson Nano™ Developer Kit (Figura 2.16) es una computadora pequeña y poderosa que permite ejecutar múltiples redes neuronales en paralelo para aplicaciones como clasificación de imágenes, detección de objetos, segmentación y procesamiento de voz. Todo en una plataforma fácil de usar que funciona con tan solo 5 vatios (NVIDIA, s.f.).

Figura 2.16: Dispositivo Jetson Nano.

2.7.2. Arduino Uno

El Arduino Uno (Figura2.17) es una placa de microcontrolador conectada a tierra en el ATmega328. Se compone de 14 pines de entrada / salida digitales (de los cuales 6 se pueden utilizar como salidas PWM), 6 entradas analógicas, un resonador cerámico de 16 MHz, una conectividad USB, un conector de alimentación, un encabezado ICSP y un botón de reinicio (N. S. Kumar et al., 2016).

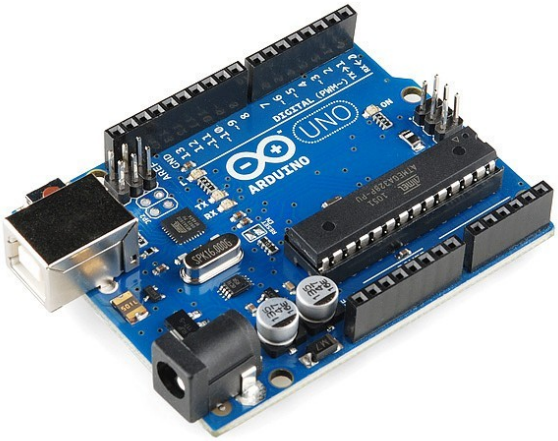


Figura 2.17: Dispositivo Arduino Uno.

2.7.3. Impresora 3D

La impresora 3D es una nueva tecnología que crea objetos físicos a partir de archivos digitales. La impresión tridimensional, también conocida como fabricación aditiva, se inventó en la década de 1980, pero ha experimentado rápidos avances en los últimos años. Desde principios del siglo XXI, las ventas de máquinas de impresión 3D han mostrado un marcado crecimiento. Hoy en día, la impresión 3D se usa ampliamente, al igual que la impresión en superficies 2D como el papel en la experiencia diaria. Ahora es posible imprimir cualquier cosa: desde pistolas hasta ropa, piezas de automóviles y joyas de diseño (Huang & X. Zhang, 2014). En la Figura 2.18 se muestra un ejemplo de una impresora 3D.

Figura 2.18: Impresora 3D.

2.8. Críterios de evaluación

La sensibilidad y la especificidad son dos medidas diferentes de un modelo de clasificación binaria. La tasa de verdaderos positivos mide la frecuencia con la que clasificamos un registro de entrada como la clase positiva y su clasificación correcta (LeCun et al., 2015). La sensibilidad cuantifica qué tan bien el modelo evita los falsos negativos (Ecuación 2.1). La

especificidad cuantifica qué tan bien el modelo evita los falsos positivos (Ecuación 2.2).

$$Sensibilidad = VP/(VP + FN) \tag{2.1}$$

$$Especificidad = VN/(VN + FP) \tag{2.2}$$

$$FFR \text{ (Tasa de rechazo erronea)} = 1 - Sensibilidad \tag{2.3}$$

$$FAR \text{ (Tasa de aceptación erronea)} = 1 - Especificidad \tag{2.4}$$

La exactitud es el grado de cercanía de las mediciones de una cantidad al valor real de esa cantidad (Ecuación 2.5)(LeCun et al., 2015).

$$Exactitud = (VP + VN)/(VP + FP + FN + VN) \tag{2.5}$$

El grado en que las mediciones repetidas en las mismas condiciones dan los mismos resultados se llama precisión en el contexto de la ciencia y la estadística. La precisión también se conoce como valor de predicción positivo (Ecuación 2.6) (LeCun et al., 2015).

$$Precisión = VP/(VP + FP) \tag{2.6}$$

En la clasificación binaria, se considera que la puntuación F1 (o puntuación F, medida F) es una medida de la precisión de un modelo. La puntuación F1 es la media armónica de las medidas de precisión y sensibilidad (descritas anteriormente) en una única puntuación (LeCun et al., 2015), se define:

$$F1 = 2VP/(2VP + FP + FN) \tag{2.7}$$

Los puntajes para F1 entre 0.0 y 1.0, donde 0.0 es el peor puntaje y 1.0 es el mejor puntaje que gustaría ver. La puntuación F1 se utiliza normalmente en la recuperación de información para ver qué tan bien un modelo recupera resultados relevantes. En el aprendizaje profundo, la puntuación F1 se utiliza como una puntuación general sobre el rendimiento de un modelo (LeCun et al., 2015).

Capítulo 3

Estado del Arte

El siguiente capítulo busca describir algunos de los trabajos existentes relacionados con el tema desarrollado en el presente trabajo. Comienza con la descripción de proyectos que tratan con cerraduras electrónicas de diferentes tipos de accionamiento ya sea de software y/o hardware.

Debido a que existe una amplia gama de aplicaciones que requieren del reconocimiento facial, se han desarrollado diferentes técnicas para alcanzar altos estándares de calidad en este ámbito. Aunado a lo anterior, empresas colosales como lo son Google y Facebook se han sumado a la investigación de nuevos desarrollos tecnológicos, por lo que con la intervención de estos dos gigantes, se han logrado grandes avances en la calidad de las redes neuronales de reconocimiento, detección y verificación facial. A continuación se describen algunos de los trabajos relacionados con este proyecto de investigación, esto con el fin de proveer una visión de lo que actualmente se está desarrollando en el mundo. Este capítulo también dispone de antecedentes del reconocimiento facial en el tema de seguridad.

3.1. Cerraduras electrónicas

En el área de cerraduras biométricas, donde características corporales de una persona son la llave de la cerradura, se encuentra el trabajo de Baidya et al., 2017, el cual propone un sistema biométrico embebido de desbloqueo de cerradura de puerta. El trabajo se basa en Arduino y un lector de huellas dactilares. La placa programable realiza la tarea de procesar la información recibida a través de los sensores, además envía señales de salida para desblo-

quear la puerta o bien emitir señales de indicación auditiva (buzzer) y visual (matriz de LED 4x4). En la Tabla 3.1 se enlistan los componentes utilizados para el desarrollo del trabajo realizado por los autores citados anteriormente.

Tabla 3.1: Lista de componentes utilizados por los autores.

Tipo de elemento	Implementado
Identificación	Huella dactilar
Lenguaje de programación	C
Procesamiento	Arduino
Sensor	Lector de huella dactilar
Audio	Buzzer
Visual	Matriz de LEDs 4x4
Interruptor	Relé
Cerradura	Cerradura electrónica

En la Tabla 3.2 se muestran los resultados obtenidos en Baidya et al., 2017 después de realizar las pruebas al sistema, en el caso I, corresponden a los intentos exitosos en estado normal y en el caso II se muestran las lecturas exitosas con pruebas realizadas con el dedo sucio o aceitoso.

Tabla 3.2: Resultados obtenidos de los intentos de lectura de huella digítal.

ID	Identidad	Intentos	Caso I	Caso II
100	Trina	20	20	18
500	Joya	20	20	19
600	Ryad	20	20	15

En Falohun et al., 2012 se realizó un proyecto de reconocimiento de iris ocular, el cual debe su funcionamiento a un sistema embebido muy práctico y efectivo según la descripción de los autores. El funcionamiento del sistema consta de una cerradura de puerta electromagnética, un escáner de iris ocular, una fuente de poder construida por los autores mismos, una PC común y un circuito integrado programable llamado 16F84, que realiza la función de intermediario entre el software y la cerradura electrónica. Este trabajo no corresponde al área de la computación, esta mayormente asociada al área de la electrónica, ya que sus descripciones principales redundan en el funcionamiento de los componentes y no tanto al sistema que hace capaz el reconocimiento del iris, pero la aportación de los autores es de suma importancia pa-

ra las instancias de hardware, ya que explican a profundidad la interacción de los dispositivos electrónicos. En la Tabla 3.3 se muestra la lista de componentes utilizados en este trabajo.

Tabla 3.3: Lista de componentes utilizados por los autores Falohun et al., 2012.

Tipo de elemento	Implementado
Identificación	Iris ocular
Lenguaje de programación	Matlab
Procesamiento	PC
Sensor	Escáner de iris ocular
Microcontrolador auxiliar	PIC16F84
Cerradura	Cerradura electromagnetica

En los trabajos de Balla & K. T. Jadhao, 2018 y Balla & K. Jadhao, 2018 desarrollaron sistemas de cerradura biométrica basados en reconocimiento facial, en ambos casos implementado en Raspberry Pi, para el primer sistema utilizaron una red neuronal convolucional llamada AlexNet y en el otro caso se utilizó un sistema de IoT que se comunica con el administrador para autorizar los accesos mediante una aplicación web y/o móvil.

Por otro lado, en el trabajo de Lwin et al., 2015, desarrollaron un sistema de cerradura biométrica basado en reconocimiento facial utilizando el método Viola-Jones, realizando la clasificación utilizando la ecuación de distancia euclidiana para identificar a la persona. A diferencia de los trabajos descritos anteriormente, este sistema está implementado en una PC convencional y en MatLab, si bien los niveles de seguridad son mucho más deficientes en comparación con otras opciones, este trabajo presume de practicidad y una buena implementación a lo pretendido por los autores.

3.2. Reconocimiento facial

El área del reconocimiento facial es muy extensa actualmente, en el trabajo de Zuo & de With, 2005 los autores desarrollaron un sistema llamado “HomeFace”. Este sistema se divide en ***n*** etapas para lograr el reconocimiento facial; (1) detección facial, (2) extracción de características e (3) identificación facial. El proceso se realiza utilizando una matriz Kernel, que detecta y extrae las características del rostro, luego se comparan con las características extraídas de los rostros de la base de datos y se busca al vecino más cercano. Según los autores

alcanzaron una efectividad del 95 % de precisión en 25 objetos de prueba, cabe mencionar que los autores no solucionan el problema de falsificación mediante fotografías impresas o digitales.

El artículo Bhattacharyya et al., 2009 presenta diferentes evaluaciones realizadas a varios sistemas de reconocimiento biométrico, entre ellos se encuentran; huella digital, facial, iris, geometría de la mano y de voz. Los autores de este estudio, utilizan 3 tipos de evaluación (por sus siglas en inglés):

- **FAR.** Tasa de aceptación erronea.
- **FRR.** Tasa de rechazo erronea.
- **EER.** Tasa de error igual ó tasa de error de cruce.

En la Tabla 3.4 se observan los resultados obtenidos por los autores en términos de las evaluaciones enlistadas anteriormente.

Tabla 3.4: Resultados obtenidos por Bhattacharyya et al., 2009.

Biometría	EER	FAR	FRR	Pruebas
Facial	NA	1 %	10 %	37437
Huella digital	2 %	2 %	2 %	25000
Geometría de la mano	1 %	2 %	2 %	129
Iris ocular	0.01 %	0.94 %	0.99 %	1224
Voz	6 %	2 %	10 %	30

El reconocimiento facial evaluado en este trabajo, presentó dificultades para la detección de rostros en las imágenes procesadas por el software. Así mismo, los autores identificaron algunos errores de identificación facial cuando se busca reconocer gemelos, uso de lentes o cambios significativos como uso de barbas, cambio de peinado, etc. También señalan que para realizar la tarea de detección de realidad se basaron en mímicas faciales, i.e., se le pide al usuario realizar movimientos como parpadear o sonreír.

Otro estudio realizado a gran escala se encuentra en Rui & Yan, 2018, en este trabajo se dedican a evaluar y clasificar múltiples sistemas de identificación biométricos. A continuación, se describen los criterios de evaluación (Tabla 3.5) y los 3 niveles de clasificación

(A-Alto, M-Medio y B-Bajo) a los cuales fueron sometidos los sistemas que se muestran en la Tabla 3.6.

En la Tabla 3.5 se describen los criterios de evaluación a los cuales se sometieron los resultados obtenidos por los autores. Estos criterios son una combinación de conceptos cualitativos y cuantitativos que en Rui & Yan, 2018 se utilizaron para evaluar una gran cantidad de sistemas similares al propuesto en el presente trabajo. El MSR es la tasa de éxito de la misión (Mission success rate), este criterio de evaluación se encarga de medir la posibilidad de resistencia a ataques de privacidad de los datos biométricos de los usuarios. Si bien, algunos pocos investigadores hacen propuestas para evitar estos ataques, no es muy común que los trabajos relacionados con este proyecto tengan estas implementaciones. Lo anterior se debe a que la privacidad de los sistemas pertenecen a otro campo muy amplio de la computación, por lo tanto la gran mayoría de las investigaciones relacionadas con el reconocimiento facial u otras características biométricas no abordan este criterio a profundidad.

Tabla 3.5: Criterios de evaluación y niveles de clasificación de los sistemas biométricos.

Criterio	Niveles		
	Alto	**Medio**	**Bajo**
FAR, FFR o EER	Los resultados experimentales de FAR, FFR o EER son menores al 3 %	Los resultados experimentales de FAR, FFR o EER estan entre el 3 % y el 10 %	Los resultados experimentales de FAR, FFR o EER son mayores al 10 %
Eficiencia	El costo de tiempo es menos de 1 segundo. O se menciona que al algoritmo solo le toma muy poco tiempo, lo cual implica que el costo computacional del método es bajo, así que es apropiado para ser implementado en un dispositivo móvil.	El costo de tiempo es de 1 a 3 segundos. O el algoritmo necesita un proceso de entrenamiento o aprendizaje, pero su requisito computacional del método es medio, esto sería posible pero no muy apropiado para un dispositivo móvil.	El costo de tiempo es más de 3 segundo. O el método este usualmente implementado en un sistema competente, y necesita un proceso de entrenamiento y aprendizaje, lo cual implica que el costo computacional de este método es alto, por lo tanto no es apropiado para un dispositivo móvil.
Universalidad	Todas las personas tienen la característica biométrica subyacente, la cual no se afecta por discapacidad, enfermedad o accidente.	Hay una pequeña probabilidad que la característica biométrica pueda ser afectada por algún accidente e.g. los mudos no pueden utilizar un sistema de autenticación de voz.	Una gran proporción de usuarios no tienen esta característica.
Unicidad	Todos los seres humanos nos diferenciamos de la característica (i.e., la característica puede únicamente representar la identidad de cada usuario y ser utilizada para autenticación).	La característica subyacente es diferente en gran escala (e.g. la probabilidad de que dos personas tengan la misma característica es menos del 0.001 %).	La característica solo es diferente en una escala pequeña (e.g. la probabilidad de dos personas tengan la misma característica es menos del 0.1 %).
Permanencia	La característica biométrica no cambia en toda la vida del usuario.	La característica no cambia distintivamente en muchos años.	La característica puede cambiar significativamente en un periodo de tiempo corto.
Aceptabilidad	De acuerdo a los resultados en Internet y nuestra propia experiencia, la carácteristica biométrica subyacente se ha usado ampliamente en autenticación en la industria y los negocios.	Una autenticación biométrica ya se ha implementado, pero no se ha usado ampliamente (i.e., el número de resultados de la búsqueda es de menos de un millón).	Hay algunos pocos ejemplos de aplicaciones prácticas.
Seguridad	Una solución de seguridad ha sido propuesta con pruebas demostradas.	La característica biométrica tiene por sí misma una particularidad de seguridad que es relativamente difícil de atacar. O se ha discutido brevemente el tema de seguridad.	Hay pocos estudios en el tema de seguridad. O la característica biométrica no es segura.
MSR	El porcentaje de MSR es mayor o igual a 90 %.	El porcentaje de MSR es mayor o igual a 50 % pero menor que 90 %.	El porcentaje de MSR es menor que 50 %.

Tabla 3.6: Resultados obtenidos por los autores.

Biometría	Referencias	Efectividad	Eficiencia	Usabilidad	Seguridad	Privacidad
Facial	González-Jiménez & Alba-Castro, 2007	B	-	M	B	-
	Queirolo et al., 2009	M	M	M	A	-
	Bhatt et al., 2012	B	-	M	-	-
	Bud, 2018	M	A	M	A	B
Iris	Pillai et al., 2011	A	-	M	M	B
	Thavalengal, Andorko et al., 2015	A	-	M	M	-
	Thavalengal, Bigioi et al., 2015	A	-	M	M	-
	Ferrer et al., 2014	A	-	M	A	-
	Baldisserra et al., 2005	A	B	M	M	B
	Li & Kot, 2010	M	-	M	M	-
	Li & Kot, 2012	A	M	M	M	-
Huella dígital de dedo o palma	Yang et al., 2014	-	-	A	M	-
	A. Kumar & Ravikanth, 2009	A	A	A	-	-
	Prasad et al., 2011	A	A	A	-	-
	Pavešić et al., 2007	A	-	M	A	-
	Pishva, 2007	-	-	M	A	-
	Jadhav & Nerkar, 2015	M	M	M	A	-
	Franco & Maltoni, 2008	M	-	M	A	-
	Ferrer et al., 2014	M	-	M	A	-
	Li & Kot, 2010	A	-	A	B	B
	Li & Kot, 2012	A	-	A	B	M
Electrocardiografía	Da Silva & Fred, 2014	-	-	M	A	-
	Carreiras et al., 2014	M	-	M	A	-
	Keshishzadeh & Rashidi, 2015	A	-	M	A	-
Voz	Jayamaha et al., 2008	B	-	A	A	-
	Gałka et al., 2014	M	-	A	B	-
	Yan & Zhao, 2016	B	-	A	A	-

3.2.1. FaceNet

En el trabajo Schroff et al., 2015 el objetivo de los autores es desarrollar un sistema llamado FaceNet, este sistema utiliza una red profunda convolucional que da tratamiento a las imágenes a procesar. El sistema es capaz de reconocer, verificar y clasificar imágenes faciales. La parte más importante o bien, donde se concentra la investigación realizada por los autores esta al final de la CNN, que es donde se clasifica, verifica o reconoce la imagen según sea el caso. El método que se utiliza para obtener los resultados se llama Triplet Loss, es una función que utiliza 3 entradas para hacer la comparación de las imágenes. La primera entrada

es la imagen base a clasificar, la segunda entrada es una imagen comparativa positiva, i.e., una imagen diferente de la misma cara para enseñar a la red neuronal que esa comparación debe ser positiva y la última entrada es una imagen comparativa negativa, i.e., una imagen de otra cara para indicar a la red neuronal que esa comparación deber ser negativa. La función Triplet Loss minimiza las distancias que existen entre la entrada base y la entrada positiva, de esa forma hace un mapeo de todas las clases en el entrenamiento. El modelo del sistema está formado principalmente por la red neuronal compuesta por diferentes arquitecturas de la red, utilizan una serie de redes pre-elaboradas pero el mismo método de clasificación de las imágenes antes mencionado. En los resultados obtenidos se muestran las tablas comparativas de efectividad de las diferentes redes utilizadas en el sistema.

3.2.2. Deep Face

Los objetivos presentados en Parkhi et al., 2015 son la creación de un dataset de gran tamaño de imágenes faciales de personas importantes del mundo, haciendo uso de servicios web gratis disponibles, se busca generar un dataset de más de dos millones de imágenes y así mismo ponerlo a disposición de la comunidad investigadora. El segundo objetivo y en el cual se enfoca mi reseña, es la investigación de algunas arquitecturas de CNN para identificación y verificación facial haciendo uso del dataset generado previamente. Los autores utilizan arquitecturas de CNN ya existentes para obtener los resultados de los tratamientos de las imágenes, estas arquitecturas vienen referenciadas y pueden ser objeto de futuras investigaciones para buscar su aplicación en implementaciones similares a este sistema, estas arquitecturas son las siguientes: Fisher Vector Faces, DeepFace, Fusion, DeepID-2,3, FaceNet y FaceNet+Alignment. Cada una de las arquitecturas con diferentes porcentajes de efectividad. Así mismo, también se realizaron pruebas con arquitecturas capaces de verificar e identificar imágenes faciales presentes en videos, para el aprendizaje de estas arquitecturas se utilizó un dataset proporcionado por YouTube llamado YouTube Faces Datasets, las arquitecturas utilizadas para las verificaciones en video son: Video Fisher Vector Faces, DeepFace, DeepID-2,2+,3 y FaceNet+Alignment. Todas las arquitecturas están enlazadas a un clasificador llamado Triplet Loss.

En el artículo Taigman et al., 2014 el objetivo esencial es lograr la verificación facial

mediante la utilización de Deep learning. Esto se busca lograr construyendo una CNN que realice los tratamientos necesarios a las imágenes. El proceso que los autores llevaron a cabo para lograr este objetivo se divide en tres partes: alineación facial, representación y métricas de verificación. La alineación facial se encarga de detectar las caras en las imágenes, este proceso se basa en el uso de puntos de referencia que detectan características correspondientes a la forma de una cara. En la etapa de representación se encuentran todos los procesos correspondientes a la extracción de las características que diferencian una cara de otra distinta. En la representación se realizan los entrenamientos a las redes neuronales profundas con el fin de que estas aprendan a distinguir cada una de las características extraídas de las imágenes previamente. Este proceso se realiza con la ayuda de una CNN que es generalmente utilizada para el tratamiento y diagnóstico de imágenes. En la etapa final del aprendizaje, las características se normalizan de 0 a 1 para reducir la sensibilidad del clasificador. En la etapa final (métricas de verificación) es donde se toman las decisiones acerca de la clasificación de las imágenes, i.e., donde se realizan las valoraciones matemáticas para determinar si una imagen pertenece o no a una clase. Aquí es donde se generan las clases o grupo de imágenes durante el entrenamiento, también aquí se representan matemáticamente los modelos que se encargaran de determinar las distancias entre una clase y otra.

3.2.3. VGG Face

Las contribuciones presentadas en Cao et al., 2018 son cuatro: la primera, crear un dataset de más de 3 millones de imágenes faciales clasificadas en más de 9000 identidades diferentes, este dataset es público y se puede utilizar libremente; segunda, se generó una línea de tiempo acerca de cómo crear un dataset con características capaces de clasificar imágenes faciales; tercera, crear una plantilla para el conjunto de prueba para explorar el rendimiento del reconocimiento de pose y edad; finalmente, se demostró que el entrenamiento de las CNN en el nuevo dataset mejoró significativamente el rendimiento en comparación con el estado del arte. Para la segunda contribución que nos describe el proceso que siguieron los autores para la creación del dataset, se realizó en 6 etapas diferentes. La primera etapa se encarga de seleccionar todos los nombres de personalidades (artistas, políticos, deportistas, etc.) disponibles en Google Image Search de los cuales cada una de las identidades elegidas debe tener al

menos 100 imágenes disponibles, de lo contrario se eliminan de la lista inicial. En la segunda etapa se realiza la búsqueda en Google Image Search de cada uno de los miembros de la lista final obtenida anteriormente, solo que en esta etapa, las búsquedas de las imágenes se realizaron con variaciones de edad y ángulos de la cara de cada uno de los personajes, obteniendo de esta forma 1400 imágenes por cada identidad. En la tercera etapa se detectan los rostros de todas las imágenes utilizando un modelo desarrollado por Zhang. En la cuarta etapa se descartan las imágenes atípicas utilizando un clasificador automático que remueve posibles imágenes erróneas. En la quinta etapa se remueven las imágenes duplicadas utilizando el clasificador VLAD. Y en la etapa final, se remueven manual y automáticamente imágenes de rostros pertenecientes a otra persona o superposiciones de rostros que no se detectaron anteriormente.

Capítulo 4

Diseño de sistema embebido

Para explicar el desarrollo del sistema embebido propuesto en este documento, primero se describe el hardware utilizado y como se lleva a cabo la comunicación entre los diferentes dispositivos utilizados para el desarrollo de esta propuesta, continuando con la descripción del software o los procesos que involucran el identificador facial.

4.1. Diseño de hardware

El sistema presentado en este trabajo, utiliza el reconocimiento facial como llave para obtener acceso a un área deseada. En la Figura 4.2 se observa el diagrama correspondiente de la relación existente entre los componentes físicos que conforman al sistema de cerradura biométrico propuesto. Los elementos de hardware que lo componen son:

- Cámara: la cámara utilizada en este proyecto es una IMX219 la cual, es un sensor de imagen de tipo de píxel activo CMOS diagonal de 4.60 mm (Tipo 1 / 4.0) con una matriz de píxeles cuadrados y 8.08M de píxeles efectivos. Este chip funciona con tres fuentes de alimentación, analógica de 2,8V, digital de 1,2V e IF de 1,8V, y tiene un bajo consumo de energía.

- Jetson Nano: La Jetson Nano ofrece 472 GFLOPS de rendimiento informático con una CPU ARM de 64 bits de cuatro núcleos y una GPU NVIDIA integrada de 128 núcleos. También incluye memoria LPDDR4 de 4 GB en un paquete eficiente de bajo consumo con modos de alimentación de 5 Watts / 10 Watts y entrada de 5V de corriente directa.

- Arduino UNO: el Arduino UNO es una placa de desarrollo con terminales para entrada o salida de datos, que se utilizan para recibir información de sensores o bien, para enviar señales de encendido/apagado.

- Cerradura electrónica: es un dispositivo fijo que al recibir una señal eléctrica su mecanismo de apertura se desbloquea para permitir la apertura de una puerta.

- Relé: es un dispositivo electrónico que tiene la función de distribuir una señal eléctrica en dos direcciones diferentes (A o B). Cuando el relé se encuentra en estado normal envía la señal en dirección A, pero cuando se activa envía la señal en dirección B.

- Fuente de alimentación: es un dispositivo electrónico que transforma la energía eléctrica de corriente alterna en corriente directa. Se utiliza comúnmente para transformar la energía eléctrica común de una casa (120 voltios de corriente alterna) en energía eléctrica para cargar dispositivos móviles (5/12/24 voltios de corriente directa).

(a) ***Arduino UNO*** **(b)** ***Relevador*** **(c)** ***Contra-chapa***

Figura 4.1: Componentes de hardware utilizados en la cerradura biométrica.

El proceso comienza en la cámara que se utiliza para dos propósitos elementales en el sistema, el primero es la toma de fotografías para agregar a los usuarios a la base de datos y el segundo para captar las imágenes de los usuarios que serán identificados facialmente.

El software del sistema es ejecutado en una microcomputadora **Jetson Nano**. Este software envía ordenes a un dispositivo **Arduino UNO** (Figura 4.1-a) que se conecta a la **Jetson Nano** vía USB para diferir las tareas de bloqueo y desbloqueo de la **cerradura electrónica**.

La **cerradura electrónica** consiste en una contra-chapa (Figura 4.1-c) que se desbloquea al recibir una señal de 12V, esta señal la recibe desde una **fuente de alimentación** de 12V, ya

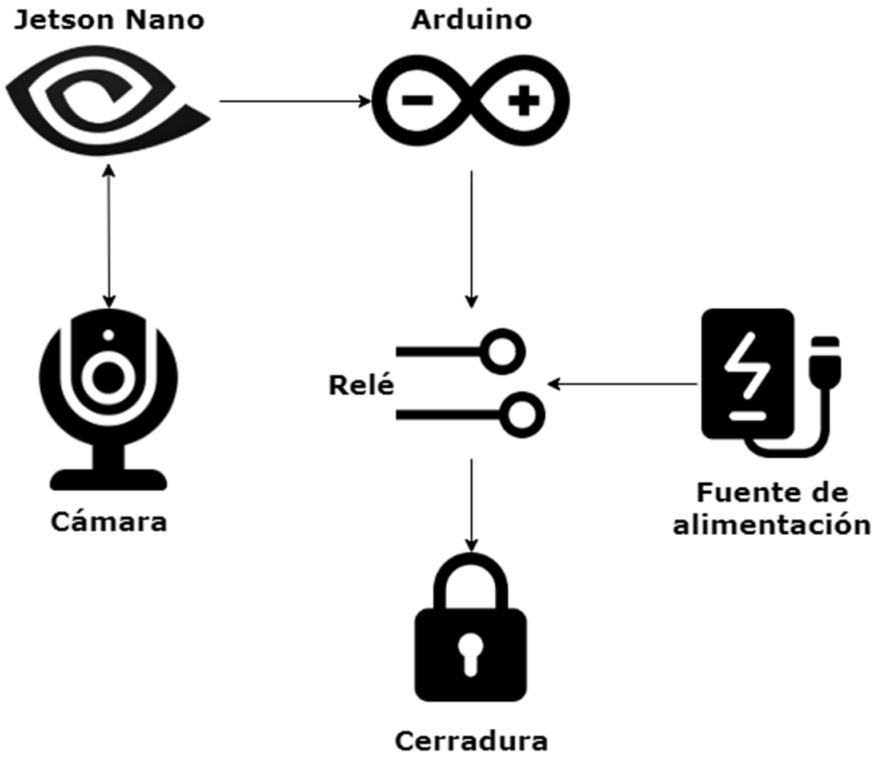


Figura 4.2: Diagrama de interacción de los dispositivos de hardware en el sistema identificador facial.

que la placa **Arduino UNO** solo es capaz de enviar voltajes máximos de 5V.

Debido a lo explicado anteriormente, es necesaria la utilización de un **relé** eléctrico (Figura 4.1-b), este dispositivo hace la función de "switch", i.e., cuando el **relé** recibe la señal (5V) del **Arduino UNO**, se activa y permite el flujo de voltaje (12V) desde la **fuente de alimentación** hasta la **cerradura electrónica**.

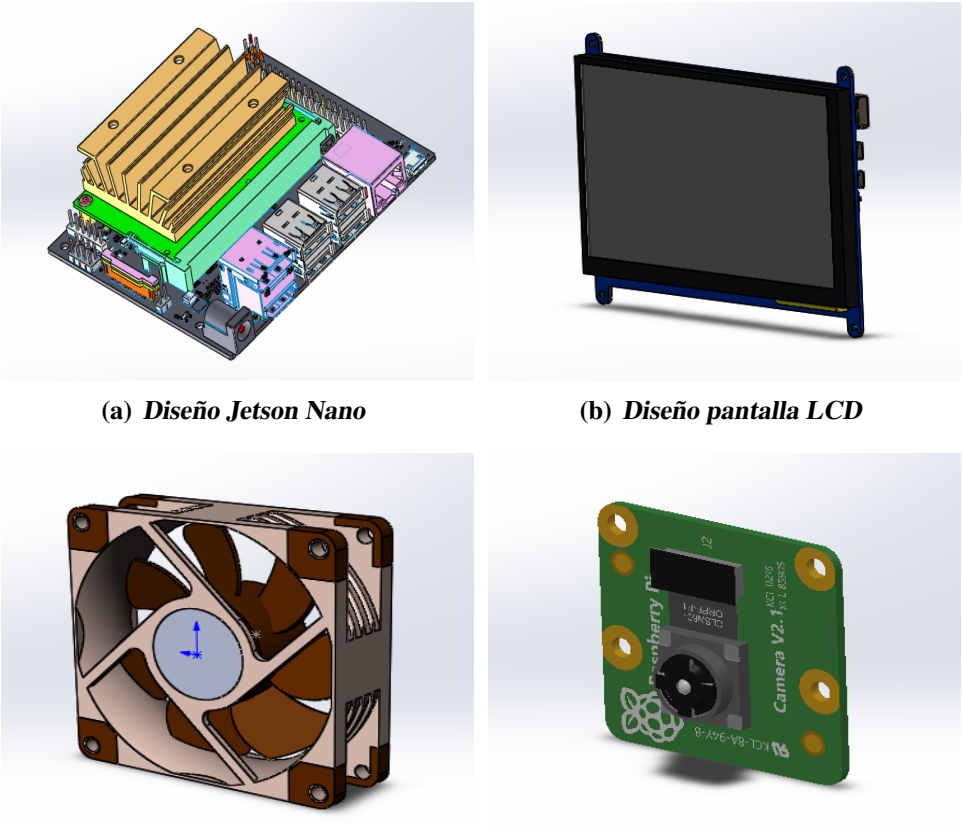
4.1.1. Diseño del prototipo

En este se explica el proceso de construcción del prototipo propuesto para este trabajo. Para realizar el diseño en 3D de los elementos que componen al prototipo se utilizó un software de diseño mecánico llamado SolidWorks. Cabe señalar que algunos de los componentes (e.g. Jetson Nano, cámara, pantalla LCD y ventilador), ya se encontraban diseñados previamente y disponibles en la página www.grabcad.com. Este sitio es una plataforma en la que millones de diseñadores comparten sus trabajos al público de forma gratuita. Los objetivos de desarrollar un diseño en un software de computadora son dos principalmente: el primero es

para corregir todos los errores que se pueden ir generando en la construcción de un prototipo, ya que, disminuye los costos de producción al visualizar virtualmente un error de cálculo que si se construyera físicamente; segundo, obtener piezas virtuales que se puedan imprimir en 3D.

4.1.1.1. Componentes electrónicos

Primero, se buscaron los diseño de los elementos principales de operación; pantalla, cámara, ventilador y Jetson Nano. Debido a que estos elementos son comúnmente utilizados para el diseño de sistemas embebidos, existen múltiples diseños disponibles en el sitio anteriormente mencionado, la Figura 4.3, se muestran los trabajos seleccionados para su uso en el prototipo de este proyecto.



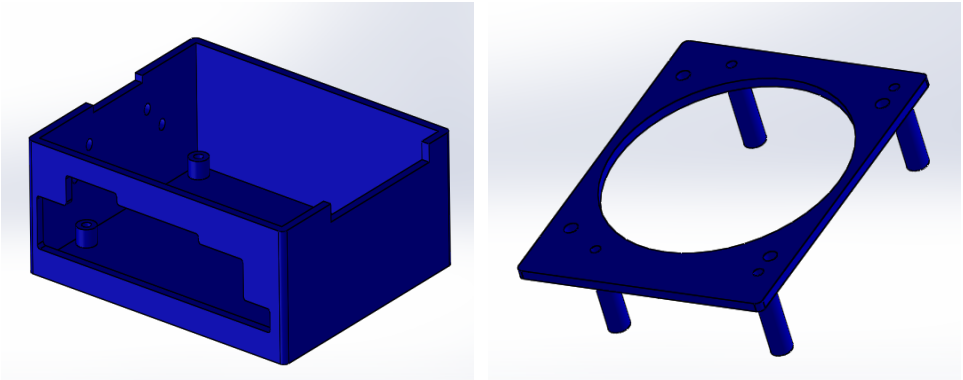
(a) ***Diseño Jetson Nano*** **(b)** ***Diseño pantalla LCD***

(c) ***Diseño ventilador*** **(d)** ***Diseño cámara***

Figura 4.3: Diseños de componentes electrónicos en SolidWorks.

4.1.1.2. Carcasa para componentes electrónicos

Después de contar con los diseños de los componentes electrónicos, se elaboraron las piezas correspondientes a la carcasa que tiene la función de proteger y contener los elementos antes mencionados. Es decir, es necesario diseñar un equivalente de gabinete para PC, pero con las características necesarias para almacenar a la Jetson Nano y el resto de elementos. El diseño de este armazón se dividió en tres partes: contenedor para Jetson Nano y ventilador, contenedor para pantalla LCD y un sujetador para la cámara. Una vez diseñadas estas piezas, se ensamblaron de manera virtual en el mismo software. En la Figura 4.4 se muestran los diseños de las piezas correspondientes al contenedor que resguarda a la Jetson Nano y en donde se monta el ventilador que se encarga de refrigerarla, este contenedor se compone de dos piezas; un pequeño cajón donde se guarda la Jetson Nano y una tapadera donde se ensambla el ventilador.

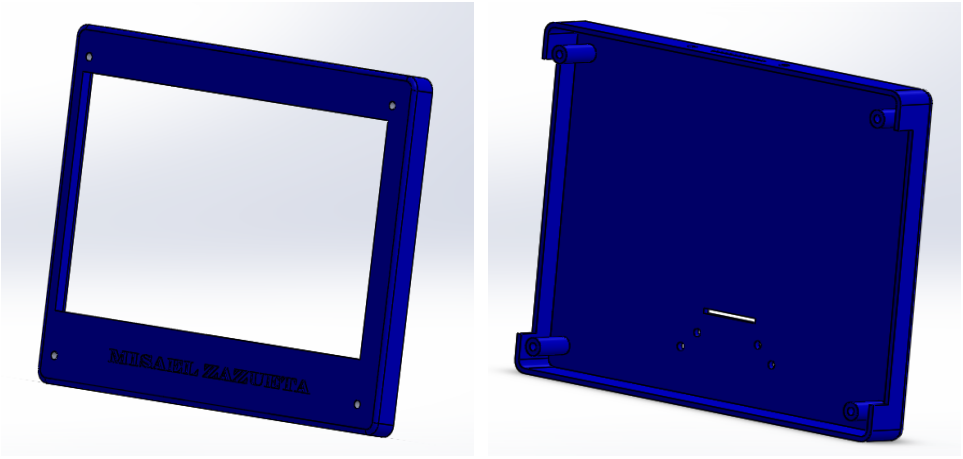


(a) ***Base para Jetson Nano*** **(b)** ***Tapadera para base de Jetson Nano***

Figura 4.4: Diseños en SolidWorks de carcasa para Jetson Nano y ventilador.

El diseño de la carcasa para proteger la pantalla LCD de 7 pulgadas, se realizó también en dos partes; una parte posterior y un marco frontal que se ensamblan con 4 tornillos colocados en las esquinas. El protector posterior, cuenta con 4 orificios en la parte baja-central para unir el marco de la pantalla con la base de la Jetson Nano. Además, en la parte superior de esta misma pieza, se colocaron dos orificios más para unir el sujetador de la cámara. En la Figura 4.5 se muestran las dos piezas correspondientes al contenedor de la pantalla LCD.

Por último, para sostener la cámara se diseñó una pieza ajustada a sus dimensiones físicas. Esta pieza (Figura 4.6) tiene un orificio en su parte inferior para permitir el paso de la cinta



(a) ***Marco frontal de pantalla LCD*** **(b)** ***Pieza posterior de pantalla LCD***

Figura 4.5: Diseños en SolidWorks de carcasa para pantalla LCD.

que comunica con la Jetson Nano.

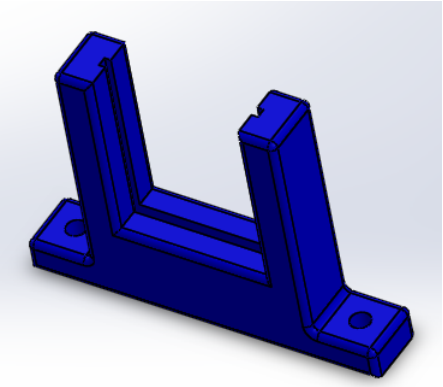
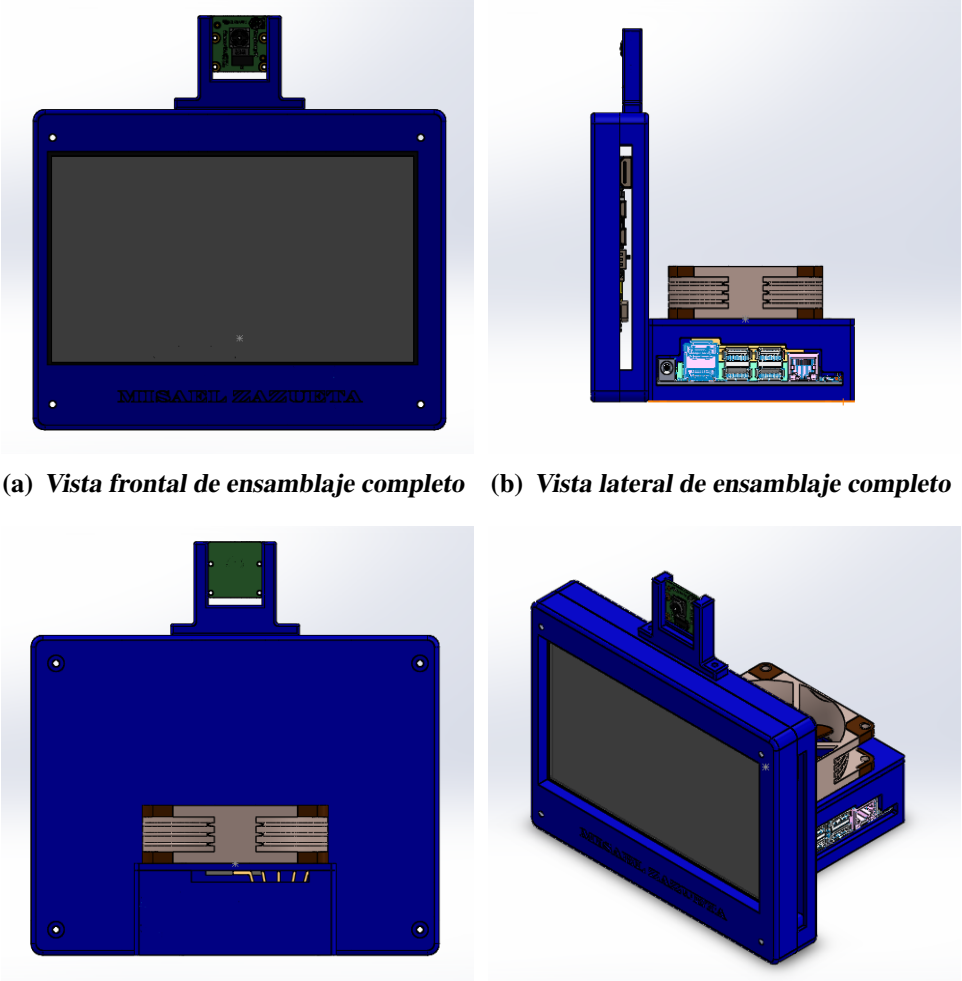


Figura 4.6: Diseño en SolidWorks de sujetador de módulo de cámara.

4.1.2. Ensamblaje de componentes electrónicos y carcasas

Una vez terminados todos los diseños, se generó una vista preliminar de lo que sería el prototipo terminado. Todos los armazones se ensamblaron para buscar errores de cálculo, de diseño, estéticos o alguna anomalía que pudiera afectar el funcionamiento de los componentes de hardware. En la Figura 4.7, se muestra el ensamblaje completo de todos los elementos que se han descrito.



(a) ***Vista frontal de ensamblaje completo*** **(b)** ***Vista lateral de ensamblaje completo***

(c) ***Vista posterior de ensamblaje completo*** **(d)** ***Vista isométrica de ensamblaje completo***

Figura 4.7: Ensamblaje de componentes electrónicos y carcasa.

4.1.3. Impresión de la carcasa

Con los diseños de las piezas completamente definidos, el siguiente paso consiste en imprimirlas con la ayuda de una impresora 3D. Este tipo de impresoras utilizan como materia prima un material polímero que toma la forma que la misma impresora le va dando según su configuración. Existen diversos materiales en el mercado que brindan diferentes características de durabilidad, resistencia, maleabilidad, etc. Para este proyecto se utilizó un material llamado PLA, que es uno de los más comunes y más utilizados por diseñadores debido a su bajo costo. En la Figura 4.8 se muestra el proceso de impresión de la base de la Jetson Nano.

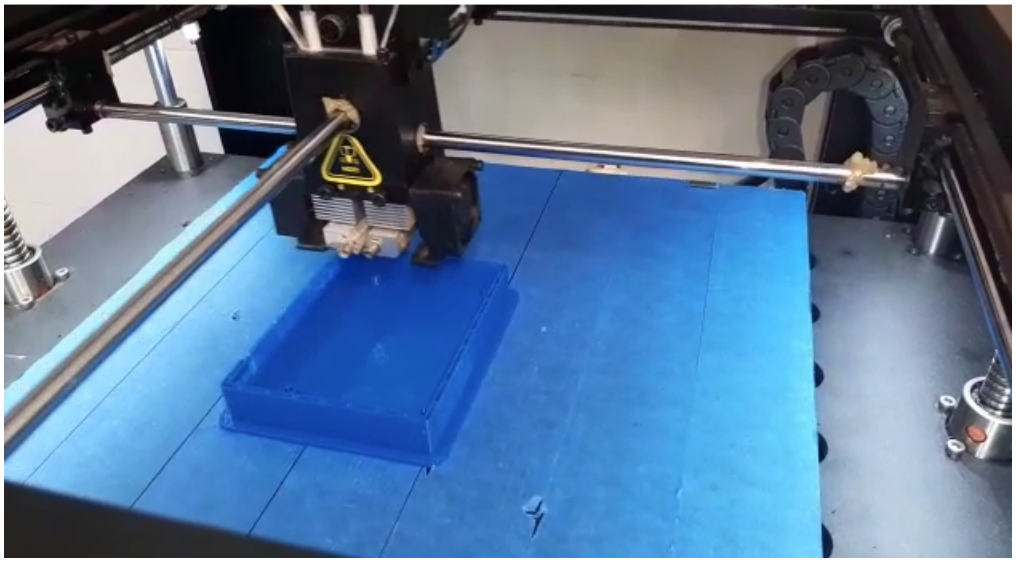


Figura 4.8: Proceso de impresión de la base de la Jetson Nano.

4.2. Desarrollo del Sistema Identificador Facial

El sistema de identificación facial propuesto tiene como entrada las imágenes recibidas y como respuesta un resultado verdadero o falso, lo que significaría el desbloqueo o bloqueo de la cerradura biométrica respectivamente.

Debido a que el software de reconocimiento facial se divide en sub-sistemas que se pueden ir incrustando una vez se terminan de desarrollar de forma individual, se tomó la decisión de utilizar una metodología iterativa e incremental que permitiera probar el sistema conforme se genere una nueva iteración. Las evaluaciones de cada una de las iteraciones fueron realizadas en conformidad de los requisitos de calidad propuestos y se conocían todas las características que comprenden al sistema, i.e., se entendían los alcances del sistema de antemano.

La posibilidad de dividir el sistema completo en partes pequeñas y el bajo riesgo de producción (pasos pequeños, riesgos pequeños)(Tan et al., 2009), fueron las razones por las cuales se utilizó la metodología iterativa e incremental.

4.2.1. Análisis de requisitos

Los requisitos del sistema se dividen en dos secciones: requisitos funcionales y requisitos no funcionales.

4.2.1.1. Requisitos no funcionales

En la Tabla 3.5 se describen los requisitos no funcionales que serán evaluados en cada iteración de desarrollo. Estos requisitos no funcionales (efectividad, eficiencia, usabilidad y seguridad) son criterios de evaluación para sistemas biométricos (Rui & Yan, 2018). Para evaluar correctamente estos parámetros, se les solicitará a los usuarios que muestren su rostro de forma frontal y cercana a la cámara.

4.2.1.2. Requisitos funcionales

El objetivo de definir los requisitos funcionales del sistema es generar una lista de operaciones que deben poderse ejecutar por el sistema para satisfacer las necesidades de los usuarios.

Los requisitos funcionales determinados para este sistema se describen de la siguiente manera:

- RF01: el sistema deberá registrar en la base de datos un directorio con el nombre de la identidad de la persona y fotografías suficientes (esto lo determina el tipo de red neuronal que se utilice) para su tratamiento dentro del directorio.
- RF02: el sistema deberá identificar facialmente a una persona que se muestre en la cámara.
- RF03: el sistema deberá mostrar el video en tiempo real en pantalla.
- RF04: el sistema deberá indicar el resultado de la identificación facial con nombre y color según el resultado sea positivo o negativo.

4.2.2. Actores

Los actores son las personas que tienen interacción con el sistema. Estos se definen de la siguiente manera:

- Administrador: se encarga de registrar a los usuarios en la base de datos del sistema.
- Usuario: son las personas que utilizaran el sistema para ser identificados facialmente.

4.2.3. Casos de uso

En los casos de uso se describen las actividades que debe hacer alguien para llevar a cabo un proceso. La lista de casos de uso son derivadas de los requisitos funcionales, se enlistan de la siguiente manera:

- CU01: Registar a un usuario
- CU02: Identificar facialmente

El CU01 indica que el administrador puede registrar a un usuario en el sistema con su nombre y fotografía. Así mismo, el CU02 indica que un usuario o el administrador pueden ser identificados facialmente

En la Figura 4.9 se observa el diagrama de interaccion de los actores con los casos de uso.

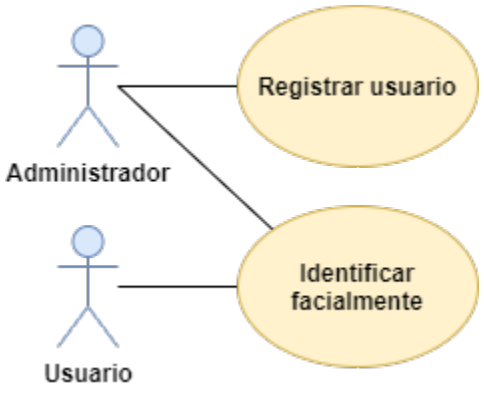


Figura 4.9: Diagrama de casos de uso

4.2.4. Diagrama de contexto

En la Figura 4.10 se muestra el diagrama de contexto del sistema. En la parte superior se muestran los actores que hace uso del sistema. En la parte media se observan los elementos de hardware que son utilizados por el sistema. Y en la parte inferior se encuentra la base de datos de la cual depende el sistema para poder identificar a los usuarios.

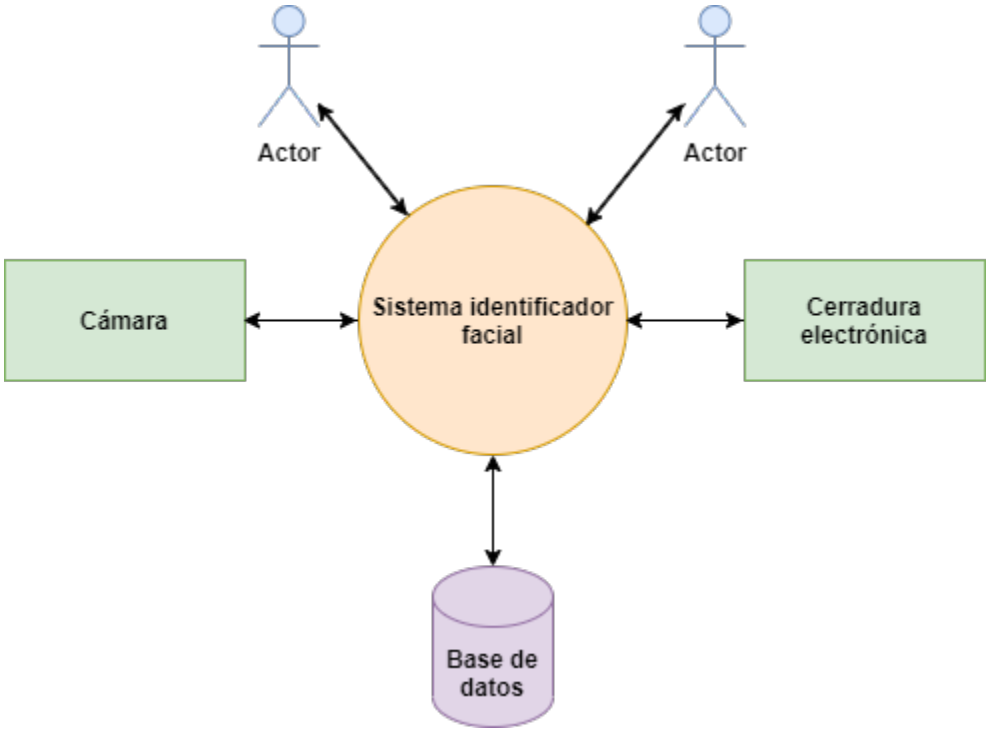


Figura 4.10: Diagrama de contexto

4.2.5. Arquetipos

Los arquetipos se definieron tomando en cuenta la interacción de los actores con el sistema para lograr cumplir los requisitos. Estos se describen de la siguiente manera:

- Usuario: es una representación del individuo que será objeto de reconocimiento.
- Identificador Facial: es una representación general del sistema de reconocimiento facial.
- Rostro: es una representación de la imagen que será tratada para el reconocimiento.

- Resultado: es una representación del resultado final del reconocimiento facial.

En la Figura 4.11 se observa la interacción que existe entre los arquetipos. Esta relación comienza cuando el Usuario provee del Rostro al Identificador Facial (imagen a tratar), una vez realizado el tratamiento de la imagen, este entrega un resultado. La razón por la cual se define por si solo al Rostro como arquetipo y no como atributo propio del Usuario, es porque cuando el Identificador Facial realiza el tratamiento de las imágenes, procede con las características propias de la imagen y no con las características que comprenden al Usuario.

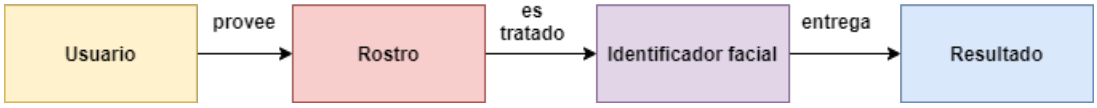


Figura 4.11: Relación de arquetipos

4.2.6. Arquitectura

El sistema identificador facial se compone de etapas que van dando tratamiento a las imágenes recibidas hasta otorgar un resultado final. Cada una de estas etapas se interpretan como filtros que extraen la información que necesitan de las imágenes. Es por esta razón que se seleccionó la arquitectura de tubos y filtros (Figura 4.12), los componentes son definidos como cajas negras y se describen de la siguiente manera:

- 1. Detección facial: recibe en su entrada una imagen y en la salida entrega un conjunto de coordenadas en dos dimensiones correspondientes a los rostros detectados en la imagen y también la misma imagen que fue recibida en la entrada.

- 2. Detección de realidad: recibe en su entrada una imagen y coordenadas de una región de la imagen y en la salida entrega las coordenadas de los rostros que correspondan a personas reales (no rostros de fotografías impresas o digitales) y la imagen recibida en la entrada con los rostros falsos señalados.

- 3. Facenet: este componente tiene dos entradas y dos salidas:

    - Entrada 1: recibe una imagen y coordenadas de una región de la imagen.

- Entrada 2: recibe una cadena de caracteres llamada etiqueta que corresponde al nombre de un usuario o la palabra “Desconocido”.
- Salida 1: entrega un vector de 128 dimensiones con las características de un rostro.
- Salida 2: entrega una imagen con los rostros etiquetados y una etiqueta de resultado (verdadero o falso).

- 4. Aumento de datos: recibe una imagen del usuario y una etiqueta (nombre) y en la salida entrega un directorio que contiene múltiples imágenes diferentes generadas a partir de la imagen original y es nombrado como la etiqueta.
- 5. Base de datos: recibe un directorio que contiene imágenes del Aumento de datos, las almacena en un repositorio local y entrega en la salida las imágenes con sus etiquetas.
- 6. Embedding: este componente tiene 3 entradas y 3 salidas:
  - Entrada 1: recibe imágenes y una etiqueta.
  - Entrada 2: recibe una imagen y coordenadas.
  - Entrada 3: recibe un vector de 128 dimensiones.
  - Salida 1: entrega una imagen.
  - Salida 2: entrega una imagen y coordenadas.
  - Salida 3: entrega una etiqueta.

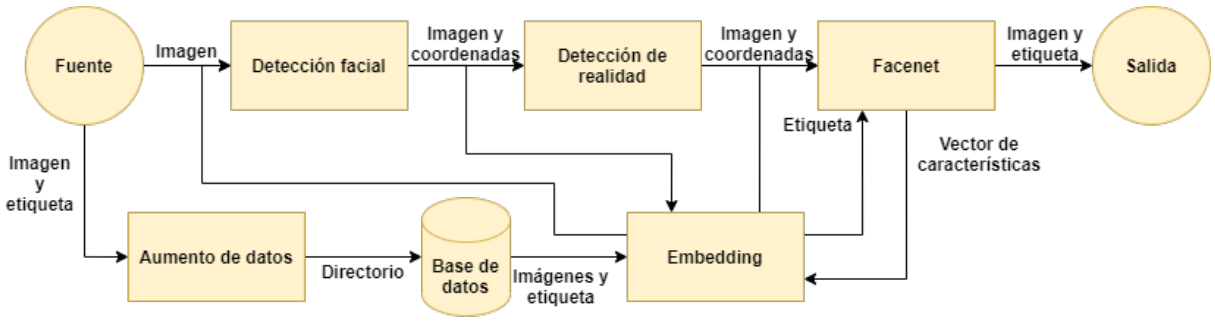


Figura 4.12: Arquitectura del sistema

4.2.7. Aplicación

El desarrollo de la aplicación se define por el flujo de los tratamientos a las imagenes, el cual se divide en dos direcciones. La primera dirección corresponde a la creación de la base de datos de usuarios que el sistema puede reconocer facialmente. La segunda realiza la tarea de reconocimiento facial en tiempo real, este bloque se puede dividir en cuatro componentes principales: (i) Detección Facial, (ii) Detección de realidad (iii), Facenet y (iv) Embedding.

4.2.7.1. Creación de la base de datos

En la Figura 4.13 se observa una extracción de los componentes de la arquitectura que tienen participación en la creación de la base de datos.

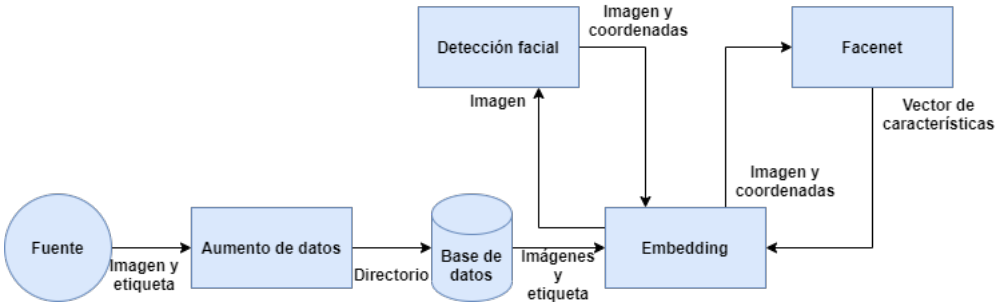


Figura 4.13: Extracción de la arquitectura para la creación de la base de datos

Para la creación de la base de datos se captura una fotografía frontal del rostro del usuario que se desea tener disponible en el sistema para su reconocimiento e introducir el nombre de la persona en la consola. El componente Aumento de datos se encarga de generar dos copias modificadas adicionales a partir de la original (Figura 4.14-a), a la primera se le aplica un ligero cambio de ángulo (Figura 4.14-b) y a la segunda un efecto de acercamiento (Figura 4.14-c). Esto con el fin de fortalecer la base de datos y mejorar el cálculo de las distancias entre las personas. El componente Embedding envía las imágenes extraídas de la base de datos al componente Detección facial, al recibir la imagen y sus coordenadas, las vuelve a enviar pero esta vez al componente Facenet, que a su vez le regresa un vector de 128 dimensiones correspondientes a las características del rostro presente en la imagen. El resultado final de este tratamiento a las imágenes disponibles en la base de datos es un archivo de extensión ".pkl", que son las distancias euclidianas de todas las identidades en un espacio de

128 dimensiones almacenado en el componente Embedding.

(a) ***Original*** (b) ***Copia 1*** (c) ***Copia 2***

Figura 4.14: Resultados del proceso de aumento de datos.

4.2.7.2. Reconocimiento facial

Este proceso requiere de la interacción de 4 componentes como se muestra en la Figura 4.15. Detección facial se encarga de detectar y extraer los rostros de una imagen. Es necesario que los usuarios que quieran ser reconocidos, muestren su rostro de manera frontal a la cámara para que la región de la imagen en donde se encuentra el rostro del usuario sea detectado, esto se logra a través de modelos o moldes en formato XML llamados Haar Cascades, los cuales contienen características de un rostro frontal en lenguaje computacional. Es decir, se realiza una comparación de estos moldes con la imagen original para encontrar los objetos que correspondan con el molde, que en este caso, se trata de un molde de un rostro frontal, si una o varias regiones de la imagen corresponden exitosamente, se extraen estas regiones en forma de coordenadas dentro de la imagen original (Figura 4.16), si una imagen contiene rostros y son detectados, se envían directamente a la siguiente etapa.

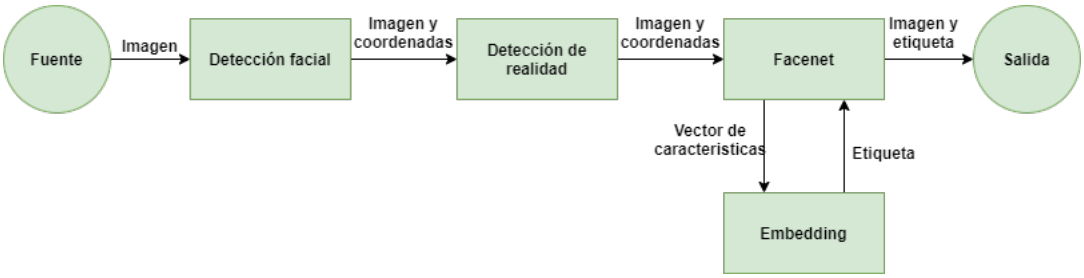


Figura 4.15: Extracción de la arquitectura para la creación de la base de datos

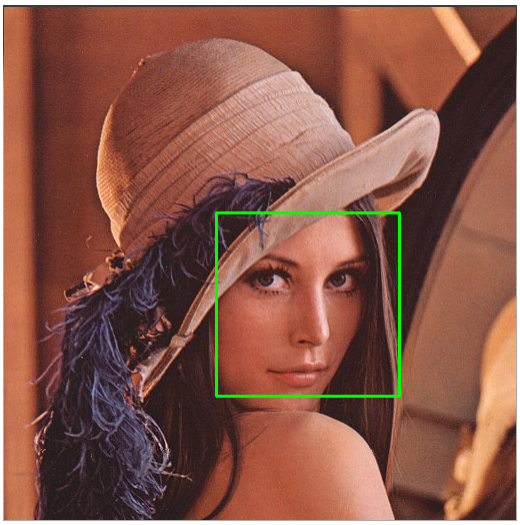
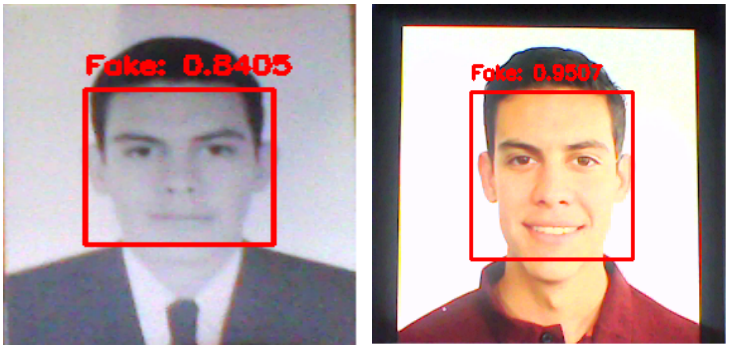


Figura 4.16: Ejemplo de una detección facial

Las CNN están compuestas por la capa de extracción de características y la capa de aprendizaje. La capa de extracción de características está formada por capas de convolución y de reducción (conocida como pooling) y la capa de aprendizaje regularmente es una red neuronal completamente conectada. Las CNN están diseñadas para procesar datos que vienen en forma de múltiples matrices, por ejemplo, una imagen en color compuesta por tres matrices de dos dimensiones que contienen intensidades de píxeles en los tres canales de color (Wolf et al., 2011). El componente Detección de realidad corresponde a la utilización de una CNN común llamada Resnet50 que se compone de 50 capas con bloques residuales (Kumaresan et al., 2021) y que clasifica los rostros extraídos en rostros reales o rostros falsos. Los rostros reales deben corresponder a los rostros de personas reales mostradas en cámara en tiempo real, mientras que los rostros falsos corresponden a intentos de engañar al sistema utilizando fotografías impresas (Figura 4.17-a) o digitales (Figura 4.17-b). Es decir, en esta etapa

se realiza un tratamiento de clasificación con una CNN a las regiones obtenidas de la etapa anterior, convirtiéndose en un filtro de imágenes de personas no reales tales como lo son las fotografías colocadas frente a la cámara con la intención de burlar al sistema.



(a) ***Fotografía impresa*** **(b)** ***Fotografía digital***

Figura 4.17: Resultados de la detección de realidad en fotografías.

Las etapas de Facenet y Embedding se realizaron con una herramienta desarrolla por Google llamada FaceNet (Schroff et al., 2015).

Por lo que, una vez generados los vectores de características durante la creación de la base de datos, las imágenes recibidas en tiempo real son procesadas de la misma manera, pero para los fines de reconocimiento, el vector generado por la imagen no se almacena, sino que se compara la distancia con cada una de las identidades disponibles en el archivo de Embedding de la base de datos y si la distancia está por debajo de un umbral, se le asigna esa identidad a la imagen (Figura 4.18).

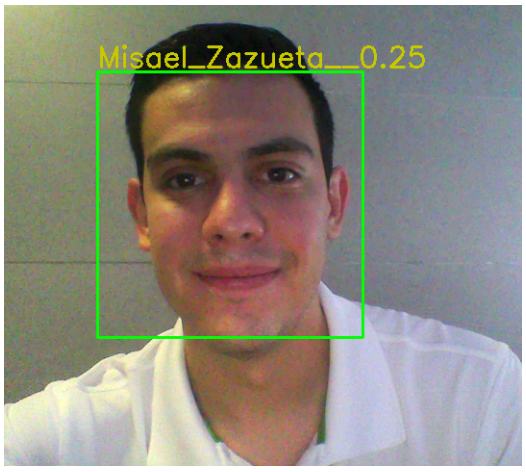


Figura 4.18: Ejemplo de un reconocimiento facial exitoso.

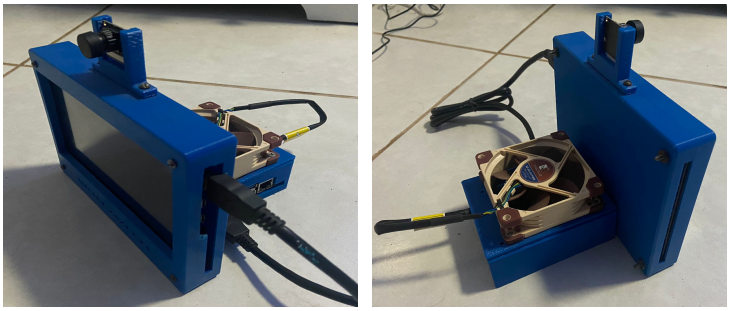
Capítulo 5

Pruebas y resultados

En este capítulo se muestran los resultados obtenidos del diseño desarrollado en SolidWorks ya impreso y ensamblado con todos sus componentes. También, se muestran los resultados de las pruebas realizadas al sistema de reconocimiento facial.

5.1. Ensamblaje de prototipo

El proceso de ensamblaje simplemente consistió en la colocación de tornillería para ajustar los elementos y evitar que se desacoplarán las piezas. En la Figura 5.1 se observa el prototipo ensamblado completamente.



(a) ***Vista frontal de prototipo*** **(b)** ***Vista posterior de prototipo***

Figura 5.1: Ensamblaje completo de prototipo.

Para probar el sistema propuesto y el prototipo diseñado, se construyó una caja de madera con un sistema de puerta accionado con los elementos de hardware descritos. El propósito de construir este simulador de puerta es para realizar pruebas del sistema en cualquier lugar. En la figura 5.2 se muestran el prototipo y el simulador de puerta construido.

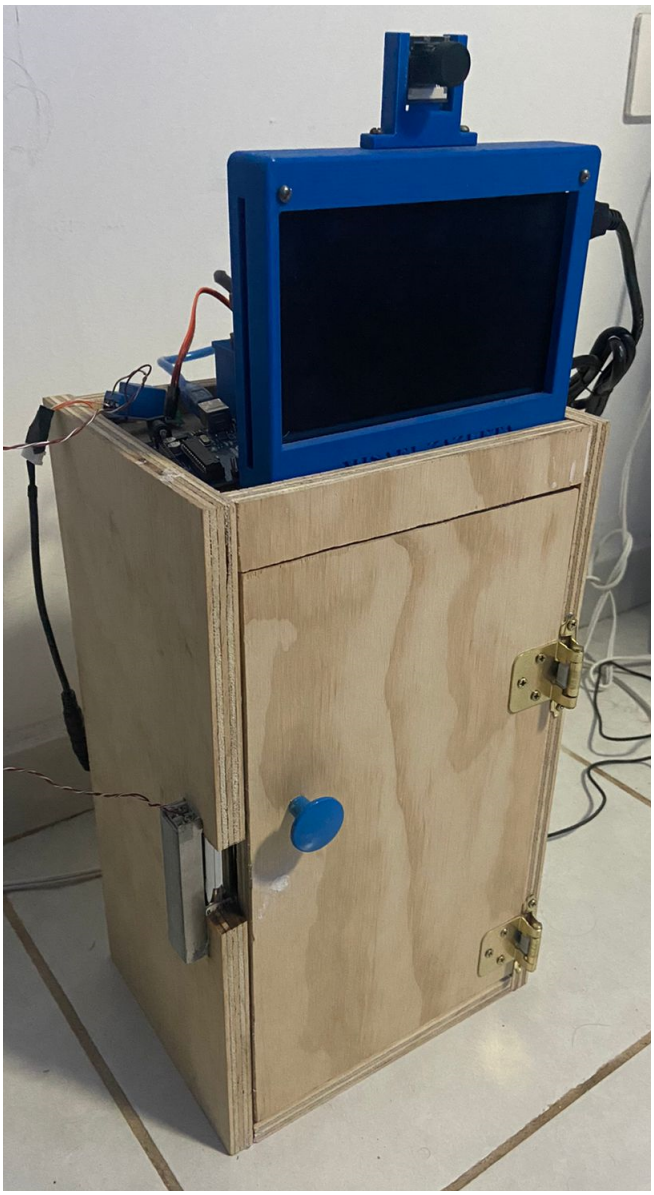


Figura 5.2: Prototipo y puerta para pruebas.

5.2. Sistema Identificador Facial

El sistema de reconocimiento facial fue desarrollado, compilado y ejecutado en lenguaje Python, únicamente una pequeña parte que se asigna a la programación del Arduino UNO fue desarrollada en lenguaje C++.

Las pruebas realizadas al sistema consistieron en la identificación de 17 usuarios presenciales, de los cuales 13 se encontraban registradas en el sistema, los resultados correspon-

dientes se encuentran en la Tabla 5.1.

Para aumentar la cantidad de pruebas de una manera virtual (Figura 5.3) se ejecutó la identificación de 40 personas virtuales, de las cuales 30 se encontraban registradas en el sistema para evaluar la correcta asignación de la identidad, con respecto a las 10 personas restantes se evaluó la asignación de la identidad "Desconocido", ya que, no se encontraban registradas en la base de datos. Las pruebas fueron realizadas con segmentos de videos de las 40 personas mencionadas anteriormente. Los videos se sometieron al identificador y se obtuvieron los resultados mostrados en la Tabla 5.2.

Tabla 5.1: Relación de identificación real y predicha por el sistema a usuarios presenciales.

Predicción \ Real	I. Correcta	Desconocido
Abigail Z.	■	
Alfonso Z.	■	
Eduardo G.	■	
Guadalupe Z.	■	
Humberto G.	■	
Humberto L.	■	
Juan B.	■	
Juan Z.	■	
Luis Z.		■
Michelle M.	■	
Mireya L.	■	
Miriam Z.		■
Misael Z.	■	
Roberto.		■
Valeria L.	■	

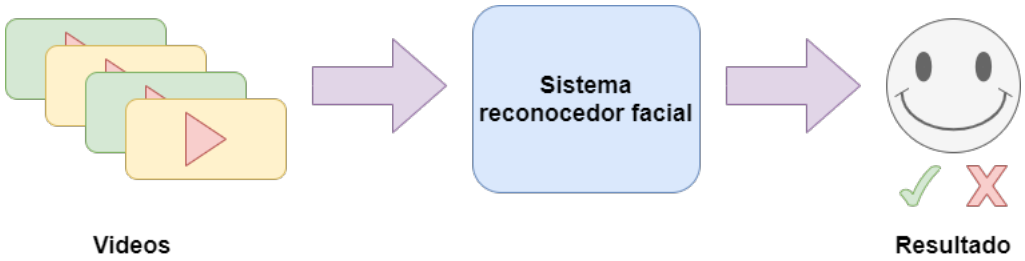


Figura 5.3: Pruebas realizadas al sistema identificador facial.

Tabla 5.2: Relación de identificación real y predicha por el sistema a usuarios virtuales.

Predicción \ Real	I. Correcta	Desconocido	Predicción \ Real	I. Correcta	Desconocido
Alexandra D.	■		Krysten R.	■	
Anya Taylor	■		Li Ka Shing		■
Asensio	■		Lionel Messi	■	
Belinda	■		Ma Huateng	■	
Bill Gates		■	Masayoshi S.	■	
Brie Larson		■	Michael Dell	■	
Carlos Vela	■		Misael Z.	■	
Casemiro	■		Morata		■
Colin Huang	■		Mukesh	■	
Cristina A.		■	Paulo D.		■
Daniel R.	■		Rhea S.	■	
Danna Paola	■		Selena G.	■	
Dietrich M.	■		Sergio Ramos	■	
Gerard P.		■	Steve B.	■	
Gonzalo H.		■	Thibaut C.	■	
Haaland	■		Tom Holland	■	
Helena B.		■	Victoria B.		■
Hui Ka Yan	■		William Lei	■	
Julia Koch	■		Yang Huiyang	■	
Karim B.		■	Zlatan I.		■

Los resultados mostrados en las Tablas 5.1 y 5.2 se indica en color verde el correcto etiquetado del rostro que corresponde a la persona en cuestión, o bien, se etiqueta de forma desconocida si esta persona no se encuentra en la base de datos del sistema. Así mismo, se muestran en rojo los errores cometidos por el identificador facial (3 errores), los cuales corresponden a personas identificadas como “Desconocido” estando registradas en la base de datos. Los aciertos del identificador facial al etiquetar de manera correcta a las personalidades disponibles en la base de datos se clasifican como “Verdaderos-Positivos”, los aciertos al no reconocer rostros que no fueron registrados se clasifican como “Verdaderos-Negativos” y los errores cometidos al no lograr identificar los rostros se clasifican como “Falsos-Negativos”.

Los resultados descritos anteriormente conciernen únicamente a pruebas realizadas con personas reales con el fin de probar el sistema completo. También se realizaron pruebas par-

ticulares a la etapa de “Detección facial”, las cuales consistieron en mostrar en cámara un total de 43 fotografías impresas y 43 fotografías digitales con el fin de obtener un etiquetado de “Fake” (falso). Como se mencionó en el capítulo 4, esta etapa funciona como filtro previo a la identificación facial, por lo que era de suma importancia realizar pruebas especializadas. De las 86 pruebas que se le practicaron a esta etapa, se obtuvieron un total de 81 fotografías identificadas exitosamente como falsas (“Verdaderos-Negativos”) y solamente 5 de ellas lograron burlar el filtro (“Falsos-Positivos”). De esta forma, los 57 rostros reales que lograron pasar este filtro de manera exitosa se clasifican como “Verdaderos-Positivos”.

Tabla 5.3: Matriz de confusión.

Predicción \ Real	Positivo	Negativo
Positivo	109	5
Negativo	3	95

La Tabla 5.3 se observa la matriz de confusión de los resultados obtenidos a las pruebas realizadas al sistema. En ella se observa 4 cuadrantes principales. En el cuadrante superior izquierdo se determinan los resultados verdaderos-positivos (VP = 109) los cuales corresponden a la correcta identificación de una persona (asignar la identidad correcta al rostro de una persona, 52 en total) o a la detección de una persona real de forma correcta (57 en total). En el cuadrante superior derecho se determinan los resultados falsos-positivos (FP = 5) los cuales corresponden a la incorrecta identificación de una persona (asignar una identidad diferente de un rostro disponible en la base de datos, 0 en total) o a detectar como persona real a una fotografía (5 en total). En el cuadrante inferior izquierdo se determinan los resultados falsos-negativos (FN = 3) los cuales corresponden la incorrecta no identificación de una persona (no asignar una identidad a un rostro disponible en la base de datos, 3 en total) o detectar como persona falsa a una persona real (0 en total). Finalmente, en el cuadrante inferior derecho se determinan los resultados verdaderos-negativos (VN = 95) los cuales corresponden a la correcta no identificación de una persona (no asignar identidad a un rostro no disponible en la base de datos, 14 en total) o bien, a identificar como persona falsa a una fotografía digital o impresa (81 en total).

En la Tabla 5.4 se encuentran los resultados alcanzados de las diferentes medidas descritas

anteriormente con base en la matriz de confusión obtenida. Estas medidas se obtuvieron realizando el cálculo descrito en las ecuaciones de los criterios de evaluación del Capítulo 2.

Tabla 5.4: Resultados obtenidos de las mediciones realizadas.

Medida	**Resultado**
Sensibilidad	0.9732
Especificidad	0.9500
Exactitud	0.9622
Precisión	0.9561
F1	0.9646
FFR	0.0268
FAR	0.0500

Si bien los resultados son ideales, esto hace referencia a los resultados obtenidos por los autores que realizaron el entrenamiento de la Facenet. Como se mencionó anteriormente la red neuronal utilizada para este proyecto alcanza un 99.63 % de rostros reconocidos exitosamente, i.e., de cada 10,000 rostros, 37 serán reconocidos de forma incorrecta, por lo que su uso para los fines establecidos en este trabajo es de extrema seguridad, ya que, por lo general la cantidad de personas que son admitidas en ciertos lugares es relativamente pequeña. Además, es necesario mencionar que la mayoría las pruebas realizadas al sistema que constaron de fragmentos de videos, favorecieran al reconocimiento, i.e., se precisó que en algún instante del vídeo el rostro de la persona se presentara de forma frontal, clara y libre de obstáculos que pudieran perjudicar el reconocimiento, además, a los usuarios presenciales se les pidió cooperación con el sistema para mostrarse de igual manera. Esto debido a que el sistema supone de un consentimiento por parte del usuario a mostrar su rostro ante la cámara para su identificación.

En la Tabla 5.5 se muestra la comparación de elementos utilizados en otros sistemas relacionados y el sistema propuesto en este trabajo.

Tabla 5.5: Comparación de componentes utilizados por otros autores.

Tipo de elemento	Baidya et al., 2017	Falohun et al., 2012	Sistema propuesto
Identificación	Huella dactilar	Iris ocular	Reconocimiento facial
Lenguaje	C	Matlab	Python
Procesamiento	Arduino	PC	Jetson Nano
Sensor	Escáner de huella d.	Escáner de iris ocular	Cámara
Audio	Buzzer		
Visual	Matriz de LEDs	Monitor de PC	Pantalla LCD
Interruptor	Relé		Relé
Microcontrolador aux		PIC16F84	Arduino UNO
Cerradura	Electrónica	Electromagnetica	Contra-chapa

En la Tabla 5.6 se muestra una comparación de resultados obtenidos en otros trabajos relacionados con este proyecto, tomando como referencia los parámetros descritos en la Tabla 3.5. El sistema propuesto demuestra un rendimiento que compite con algunos de los existentes en el área, desafortunadamente la gran mayoría de los investigadores que realizan este tipo de proyectos únicamente analizan los resultados obtenidos por el rendimiento del hardware y no así del software, por lo que se dificulta una comparación natural de sistemas similares.

Tabla 5.6: Comparación de otros sistemas con el sistema propuesto en este trabajo.

Biometría	Referencias	Efectividad	Eficiencia	Usabilidad	Seguridad	Privacidad
Facial	González-Jiménez & Alba-Castro, 2007	B	-	M	B	-
	Queirolo et al., 2009	M	M	M	A	-
	Bhatt et al., 2012	B	-	M	-	-
	Bud, 2018	M	A	M	A	B
	Propuesta	A	A	M	A	-

Capítulo 6

Conclusiones

A continuación se describen las conclusiones, aportaciones y trabajos a futuro, con base en todo lo presentado en los capítulos anteriores, se muestra la relevancia que puede tener este trabajo, los beneficios y las mejoras que se pueden realizar en el porvenir de la investigación desde la opinión del autor.

6.1. Conclusiones

La seguridad habitacional es una necesidad que ha ido creciendo conforme aumenta el valor de los bienes adquiridos por las personas, aunque actualmente existen diversas tecnologías capaces de brindar altos estándares de seguridad y vigilancia residencial, estas suelen ser de altos costos y resultan inasequibles para la mayoría de la población.

Los técnicos e ingenieros relacionados con el área de la domótica, han hecho grandes esfuerzos por generar más y mejores herramientas de seguridad para los hogares. Una parte importante de estos avances es gracias al uso de técnicas de aprendizaje profundo, tales como la utilización de redes neuronales que aprenden patrones de comportamiento de los seres humanos para facilitar actividades diarias, p. ej., encender o apagar aparatos electrónicos que representan algún peligro, alarmas, aparatos de jardinería, seguridad, etc.

El interés en sistemas automáticos ha crecido a lo largo de la historia y actualmente ese crecimiento sigue en auge, el ser humano ha desarrollado tecnologías que disminuyen el esfuerzo humano o bien, que provean mayor seguridad en tareas difíciles, peligrosas o que requieran de altos niveles de precisión. Es por eso que el desarrollo de técnicas diferentes

promete el avance tecnológico y por ende, mejores productos y a menor precio de mercado. La instalación de un sistema como el propuesto en el trabajo presente, podría reducir en más de un 50 % del costo respecto a los sistemas presentes en el mercado actual.

Este trabajo brinda la oportunidad de generar sistemas de cerradura electrónica de buena calidad y de bajo costo. Si bien, la diferencia de precios del presente proyecto y las opciones actuales del mercado puede ser muy poca, existe la oportunidad de beneficiarse de manera exponencial con el uso de técnicas de Internet de las Cosas, i.e., un solo prototipo podría funcionar para múltiples accesos de una edificación, disminuyendo los costos de esa forma. El resultado final es un prototipo de dimensiones pequeñas capaz de ejecutar el sistema de reconocimiento facial con una velocidad de respuesta competitible, cabe la posibilidad también de realizar cambios relacionados con el tipo de cerradura a utilizar debido a que los módulos electrónicos (véase Figura 4.1) seleccionados son fácilmente permutables.

Las pruebas descritas en el capítulo anterior demuestran que el sistema tiene excelentes grados en los diferentes modelos de evaluación realizados (todos por encima del 95 %), incluso si se utiliza un alto número de usuarios registrados. En su conjunto (hardware y software) el sistema propuesto, ofrece una notable opción para quienes requieren de un sistema de cerradura.

6.2. Aportaciones

La aportación más importante de este trabajo es el sistema de reconocimiento facial. Obtiene altos estándares de las mediciones realizadas, lo cual significa que se trata de un sistema viable para su utilización en sistemas de seguridad con bajo costo de recursos. A continuación se enlistan las aportaciones más relevantes:

- Diseño de circuito de potencia (Arduino UNO - relé - cerradura) mantenible y funcional.
- Diseño de sistema embebido completo para la ejecución del software y apertura de puerta.
- Utilización de microcomputadora Jetson Nano para ejecución de técnicas de aprendi-

zaje profundo.

- Desarrollo de sistema de identificación facial utilizando FaceNet y una red neuronal convolucional.
- Introducir etapa de detección de realidad previo al reconocimiento facial.
- Altos grados obtenidos en los criterios de evaluación de clasificación en técnicas de aprendizaje profundo.

6.3. Trabajo Futuro

A lo largo del desarrollo de este trabajo se visualizaron ideas prometedoras para el futuro. Un ejemplo de estas ideas, se trata de incorporar tecnologías de Internet de las Cosas para lograr un sistema de seguridad habitacional completo, controlando todos los accesos desde un mando de control. Otra vía de trabajo es no depender únicamente del reconocimiento facial como herramienta de identidad, sino también desarrollar técnicas de identificación por voz, por huella digital y/o iris ocular, esto se traduciría en un sinfín de combinaciones disponibles para aumentar la seguridad y versatilidad de los sistemas. Por último, en el tema de la detección de realidad es viable utilizar un sistema distinto a la CNN, con el fin de aumentar la seguridad. Una opción es la utilización de cámaras infrarrojas capaces de obtener lecturas de calor.

Bibliografía

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., ... Zheng, X. (2016). TensorFlow: A System for Large-Scale Machine Learning. *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 265-283. https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi (page 22)

Adeli, H., & Yeh, C. (1989). Perceptron learning in engineering design. *Computer-Aided Civil and Infrastructure Engineering*, *4*(4), 247-256 (page 16).

Asimov, I. (2004). *I, robot* (Vol. 1). Spectra. (Page 11).

Baidya, J., Saha, T., Moyashir, R., & Palit, R. (2017). Design and implementation of a fingerprint based lock system for shared access. *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, 1-6. https://doi.org/10.1109/CCWC.2017.7868448 (pages 28, 29, 61)

Baldisserra, D., Franco, A., Maio, D., & Maltoni, D. (2005). Fake Fingerprint Detection by Odor Analysis. En D. Zhang & A. K. Jain (Eds.), *Advances in Biometrics* (pp. 265-272). Springer Berlin Heidelberg. (Page 34).

Balla, P. B., & Jadhao, K. T. [K. T.]. (2018). IoT Based Facial Recognition Security System. *2018 International Conference on Smart City and Emerging Technology (ICSCET)*, 1-4. https://doi.org/10.1109/ICSCET.2018.8537344 (page 30)

Balla, P. B., & Jadhao, K. (2018). IoT based facial recognition security system. *2018 International Conference on Smart City and Emerging Technology (ICSCET)*, 1-4 (page 30).

Bhatt, H. S., Bharadwaj, S., Singh, R., & Vatsa, M. (2012). Recognizing surgically altered face images using multiobjective evolutionary algorithm. *IEEE Transactions on Information Forensics and Security*, *8*(1), 89-100 (pages 34, 61).

Bhattacharyya, D., Ranjan, R., Alisherov, F., Choi, M., et al. (2009). Biometric authentication: A review. *International Journal of u-and e-Service, Science and Technology*, *2*(3), 13-28 (page 31).

Buchanan, B. G. (2005). A (very) brief history of artificial intelligence. *Ai Magazine*, *26*(4), 53-53 (page 12).

Bud, A. (2018). Facing the future: The impact of Apple FaceID. *Biometric technology today*, *2018*(1), 5-7 (pages 34, 61).

Cao, Q., Shen, L., Xie, W., Parkhi, O. M., & Zisserman, A. (2018). VGGFace2: A Dataset for Recognising Faces across Pose and Age. *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*, 67-74. https://doi.org/10.1109/FG.2018.00020 (page 36)

Carreiras, C., Lourenço, A., Fred, A., & Ferreira, R. (2014). ECG signals for biometric applications - are we there yet? *2014 11th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, *02*, 765-772. https://doi.org/10.5220/0005160507650772 (page 34)

Coppin, B. (2004). *Artificial intelligence illuminated*. Jones & Bartlett Learning. (Page 12).

Coşkun, M., Uçar, A., Yildirim, Ö., & Demir, Y. (2017). Face recognition based on convolutional neural network. *2017 International Conference on Modern Electrical and Energy Systems (MEES)*, 376-379. https://doi.org/10.1109/MEES.2017.8248937 (page 19)

Da Silva, H. P., & Fred, A. (2014). Harnessing the power of biosignals. *Computer*, *47*(03), 74-77 (page 34).

Divya, R. S., & Mathew, M. (2017). Survey on various door lock access control mechanisms. *2017 International Conference on Circuit ,Power and Computing Technologies (ICCPCT)*, 1-3. https://doi.org/10.1109/ICCPCT.2017.8074187 (pages 6, 8-10)

Falohun, A., Omidiora, E., Fakolujo, O., Afolabi, O., Oke, A., & Ajala, F. (2012). Development of a biometrically-controlled door system (using iris), with power backup. *American Journal of Scientific and Industrial Research*, *3*(4), 203-207 (pages 29, 30, 61).

Ferrer, M. A., Morales, A., & Dıaz, A. (2014). An approach to SWIR hyperspectral hand biometrics. *Information Sciences*, *268*, 3-19 (page 34).

Fezari, M., & Al Dahoud, A. (2018). Integrated Development Environment "IDE" For Arduino. *WSN applications*, 1-12 (page 23).

Franco, A., & Maltoni, D. (2008). Fingerprint synthesis and spoof detection. En *Advances in Biometrics* (pp. 385-406). Springer. (Page 34).

Gałka, J., Masior, M., & Salasa, M. (2014). Voice authentication embedded solution for secured access control. *IEEE Transactions on Consumer Electronics*, *60*(4), 653-661 (page 34).

García, V. H., & Vega, N. (2018). Low Power Sensor Node Applied to Domotic Using IoT. En M. F. Mata-Rivera & R. Zagal-Flores (Eds.), *Telematics and Computing* (pp. 56-69). Springer International Publishing. (Page 1).

González-Jiménez, D., & Alba-Castro, J. L. (2007). 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 (pages 34, 61).

Huang, W., & Zhang, X. (2014). 3D printing: print the future of ophthalmology. *Investigative ophthalmology & visual science*, *55*(8), 5380-5381 (page 26).

INEGI. (2021). Incidencia delictiva [[Web; accedido el 01-02-2021]]. https://www.inegi.org.mx/temas/incidencia/. (Page 1)

Jadhav, M., & Nerkar, P. M. (2015). Implementation of an embedded hardware of FVRS on FPGA. *2015 International Conference on Information Processing (ICIP)*, 48-53. https://doi.org/10.1109/INFOP.2015.7489349 (page 34)

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, *29*(3), 31-44 (page 16).

Jayamaha, R. G. M. M., Senadheera, M. R. R., Gamage, T. N. C., Weerasekara, K. D. P. B., Dissanayaka, G. A., & Kodagoda, G. N. (2008). VoizLock - Human Voice Authentication System using Hidden Markov Model. *2008 4th International Conference on Information and Automation for Sustainability*, 330-335. https://doi.org/10.1109/ICIAFS.2008.4783977 (page 34)

Jose, E., M., G., Haridas, M. T. P., & Supriya, M. (2019). Face Recognition based Surveillance System Using FaceNet and MTCNN on Jetson TX2. *2019 5th International*

Conference on Advanced Computing Communication Systems (ICACCS), 608-613. https://doi.org/10.1109/ICACCS.2019.8728466 (page 19)

Keshishzadeh, S., & Rashidi, S. (2015). Single lead Electrocardiogram feature extraction for the human verification. *2015 5th International Conference on Computer and Knowledge Engineering (ICCKE)*, 118-122. https://doi.org/10.1109/ICCKE.2015.7365870 (page 34)

Kumar, A., & Ravikanth, C. (2009). Personal authentication using finger knuckle surface. *IEEE Transactions on Information Forensics and Security*, *4*(1), 98-110 (page 34).

Kumar, N. S., Vuayalakshmi, B., Prarthana, R. J., & Shankar, A. (2016). IOT based smart garbage alert system using Arduino UNO. *2016 IEEE Region 10 Conference (TENCON)*, 1028-1034. https://doi.org/10.1109/TENCON.2016.7848162 (page 25)

Kumaresan, S., Aultrin, K. S. J., Kumar, S. S., & Anand, M. D. (2021). Transfer Learning With CNN for Classification of Weld Defect. *IEEE Access*, *9*, 95097-95108. https://doi.org/10.1109/ACCESS.2021.3093487 (page 53)

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444 (pages 14, 15, 18, 26, 27).

Li, S., & Kot, A. C. (2012). Fingerprint combination for privacy protection. *IEEE transactions on information forensics and security*, *8*(2), 350-360 (page 34).

Li, S., & Kot, A. C. (2010). Privacy protection of fingerprint database. *IEEE Signal Processing Letters*, *18*(2), 115-118 (page 34).

Lwin, H. H., Khaing, A. S., & Tun, H. M. (2015). Automatic door access system using face recognition. *International Journal of scientific & technology research*, *4*(06), 294-99 (page 30).

McCarthy, J. (2007). What is artificial intelligence. *Computer Science Department, Stanford University* (page 12).

McCauley, L. (2007). AI armageddon and the three laws of robotics. *Ethics and Information Technology*, *9*(2), 153-164 (page 11).

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of machine learning*. MIT press. (Pages 12, 13).

Nilsson, N. J. (1996). Artificial intelligence: A modern approach: Stuart Russell and Peter Norvig, (Prentice Hall, Englewood Cliffs, NJ, 1995); xxviii + 932 pages. *Artificial Intelligence*, *82*(1), 369-380. https://doi.org/https://doi.org/10.1016/0004-3702(96)00007-0 (page 12)

NVIDIA. (s.f.). Jetson Nano Developer Kit [[Web; accedido el 29-01-2021]]. https://developer.nvidia.com/embedded/jetson-nano-developer-kit. (Page 24)

Oke Alice, O., Adigun Adebisi, A., Falohun Adeleye, S., & Alamu, F. (2013). Development of a programmable electronic digital code lock system. *International Journal of Computer and Information Technology (ISSN: 2279–0764) Volume* (page 7).

Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition (page 35).

Patterson, J., & Gibson, A. (2017). *Deep learning: A practitioner's approach*. O'Reilly Media, Inc. (Page 13).

Pavešić, N., Savič, T., Ribarić, S., & Fratrić, I. (2007). A multimodal hand-based verification system with an aliveness-detection module. *Annales des télécommunications*, *62*(1), 130-155 (page 34).

Pavithra, D., & Balakrishnan, R. (2015). IoT based monitoring and control system for home automation. *2015 Global Conference on Communication Technologies (GCCT)*, 169-173. https://doi.org/10.1109/GCCT.2015.7342646 (page 10)

Pillai, J. K., Patel, V. M., Chellappa, R., & Ratha, N. K. (2011). Secure and robust iris recognition using random projections and sparse representations. *IEEE transactions on pattern analysis and machine intelligence*, *33*(9), 1877-1893 (page 34).

Pishva, D. (2007). Spectroscopic Approach for Aliveness Detection in Biometrics Authentication. *2007 41st Annual IEEE International Carnahan Conference on Security Technology*, 133-137. https://doi.org/10.1109/CCST.2007.4373480 (page 34)

Prasad, S., Govindan, V., & Sathidevi, P. (2011). Palmprint authentication using fusion of wavelet and contourlet features. *Security and Communication Networks*, *4*(5), 577-590 (page 34).

Queirolo, C. C., Silva, L., Bellon, O. R., & Segundo, M. P. (2009). 3D face recognition using simulated annealing and the surface interpenetration measure. *IEEE transactions on pattern analysis and machine intelligence*, *32*(2), 206-219 (pages 34, 61).

Reddy, E. J., Sridhar, C., & Rangadu, V. P. (2016). Research and development of knowledge based intelligent design system for bearings library construction using SolidWorks API. En *Intelligent Systems Technologies and Applications* (pp. 311-319). Springer. (Page 24).

Rui, Z., & Yan, Z. (2018). A survey on biometric authentication: Toward secure and privacy-preserving identification. *IEEE access*, *7*, 5994-6009 (pages 31, 32, 46).

Sanner, M. F., et al. (1999). Python: a programming language for software integration and development. *J Mol Graph Model*, *17*(1), 57-61 (page 22).

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 815-823 (pages 19-21, 34, 54).

Sharma, N., Jain, V., & Mishra, A. (2018). An analysis of convolutional neural networks for image classification. *Procedia computer science*, *132*, 377-384 (page 18).

Shirai, Y. (2012). *Three-dimensional computer vision*. Springer Science & Business Media. (Page 19).

Srinivasan, D., Liew, A., & Chang, C. (1994). A neural network short-term load forecaster. *Electric Power Systems Research*, *28*(3), 227-234 (page 17).

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going Deeper With Convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (page 20).

Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1701-1708 (page 35).

Tan, T., Li, Q., Boehm, B., Yang, Y., He, M., & Moazeni, R. (2009). Productivity trends in incremental and iterative software development. *2009 3rd International Symposium on Empirical Software Engineering and Measurement*, 1-10. https://doi.org/10.1109/ESEM.2009.5316044 (page 45)

Thavalengal, S., Andorko, I., Drimbarean, A., Bigioi, P., & Corcoran, P. (2015). 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 (page 34).

Thavalengal, S., Bigioi, P., & Corcoran, P. (2015). Iris authentication in handheld devices-considerations for constraint-free acquisition. *IEEE Transactions on Consumer Electronics*, *61*(2), 245-253 (page 34).

Turing, A. (1950). Computing Machinery and Intelligence. *Mind*, *59*(October), 433-60. https://doi.org/10.1093/mind/LIX.236.433 (page 12)

Verma, G. K., & Tripathi, P. (2010). A digital security system with door lock system using RFID technology. *International Journal of Computer Applications*, *5*(11), 6-8 (page 8).

William, I., Rachmawanto, E. H., Santoso, H. A., Sari, C. A., et al. (2019). Face recognition using FaceNet (survey, performance test, and comparison). *2019 Fourth International Conference on Informatics and Computing (ICIC)*, 1-6 (pages 21, 23).

Wolf, L., Hassner, T., & Maoz, I. (2011). Face recognition in unconstrained videos with matched background similarity. *CVPR 2011*, 529-534 (page 53).

Xu, X., Du, M., Guo, H., Chang, J., & Zhao, X. (2020). Lightweight FaceNet Based on MobileNet. *International Journal of Intelligence Science*, *11*(1), 1-16. https://doi.org/10.4236/ijis.2021.111001 (page 20)

Yan, Z., & Zhao, S. (2016). A usable authentication system based on personal voice challenge. *2016 International Conference on Advanced Cloud and Big Data (CBD)*, 194-199 (page 34).

Yang, W., Hu, J., & Wang, S. (2014). 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 (page 34).

Zhang, X.-D. (2020). Machine learning. En *A Matrix Algebra Approach to Artificial Intelligence* (pp. 223-440). Springer. (Page 13).

Zou, J., Han, Y., & So, S.-S. (2008). Overview of artificial neural networks. *Artificial Neural Networks*, 14-22 (page 15).

Zuo, F., & de With, P. (2005). Real-time embedded face recognition for smart home. *IEEE transactions on consumer Electronics*, *51*(1), 183-190 (page 30).