ISSN 1870-4069

Sistema de control de acceso mediante identificación facial usando aprendizaje profundo

José Misael Burruel Zazueta[1], Hector Rodríguez Rangel[1], Gloria Ekaterine Peralta Peñuñuri[1], Víctor Alejandro González Huitrón[1], Luis Alberto Morales Rosales[2]

[1] Tecnológico Nacional de México, División de Posgrado, Maestría en Ciencias de la Computación, México

[2] Universidad Michoacana de San Nicolás de Hidalgo, Consejo Nacional de Ciencia y Tecnología, México

{jose_burruel, hrodriguez, gperalta, victor.gonzalez}@itculiacan.edu.mx, lamorales@conacyt.mx

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. En este articulo, se propone un método de identificación facial para su implementación en sistemas de cerraduras biométricos. Este método esta 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 documentados, el sistema de identificación facial logra más del 99 % de tasa de reconocimiento correcto y un alto rendimiento.

Palabras clave: Cerradura biométrica, aprendizaje profundo, reconocimiento facial, red neuronal convolucional, facenet.

Access Control System through Facial Identification Using Deep Learning

Abstract. There are various technologies used for the automatic unlocking of doors, some of these techniques use biometric systems to analyze the user's body and thus ensure that he is a desirable person. In this article, a facial identification method is proposed for its implementation in biometric lock systems. This method is mainly based on the use of a convolutional neural network developed by Google called Facenet. Likewise, an embedded electronic lock system is

proposed to be implemented through the facial identifier. Different from other documented projects, the facial identification system achieves more than 99% correct recognition rate and high performance.

Keywords: Biometric lock, deep learning, facial recognition, convolutional neural network, facenet.

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 [6]. Poca vigilancia, sistemas de seguridad deficientes y otros factores, han contribuido a que los robos en propiedades privadas tengan tasas de incidencia alarmantes. La seguridad representa la protección de nuestra vida y activos. Garantizar la seguridad de las personas y sus cosas de valor es muy importante para prevenir la manipulación ilegal.

Por lo tanto, centrarse principalmente en la seguridad de la cerradura de la puerta o la seguridad de la puerta es muy importante para evitar problemas adicionales en el área monitorizada [10]. 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 especificas. En [5] se mencionan algunas tecnologías empleadas en la seguridad de las cerraduras de puertas (e.g. mecánica, biométrica, por contraseña, entre otras).

El aprendizaje profundo permite que los modelos computacionales que se componen de múltiples capas de procesamiento, aprendan representaciones de datos con múltiples niveles de abstracción. Estos métodos han mejorado de forma drástica el estado de la técnica en reconocimiento de voz, reconocimiento de objetos visuales, detección de objetos y muchos otros dominios [16].

El uso del aprendizaje profundo ha generado múltiples técnicas que se han popularizado exponencialmente en la ultima década debido a los altos estándares de calidad alcanzados. La seguridad habitacional no es la excepción, ya que se han realizado múltiples esfuerzos por mejorar los sistemas de seguridad basados en lecturas biométricas mediante aprendizaje profundo.

La principal motivación del presente trabajo es desarrollar un sistema de cerradura con altos estándares de calidad y con una mínima interacción con el usuario. Se busca reducir la relación del ser humano con el sistema para evitar errores por parte de los usuarios, el ejemplo mas común es extraviar objetos de manipulación de las cerraduras como llaves o tarjetas electrónicas, esto también supondría un problema de seguridad mayor si una persona no autorizada obtiene estos objetos.

El proyecto en general propone un sistema de cerradura biométrico basado en identificación facial. La realización del sistema de cerradura biométrico comprende diferentes procesos como son:

1. La identificación facial, donde el sistema identifica al sujeto a partir de su rostro.

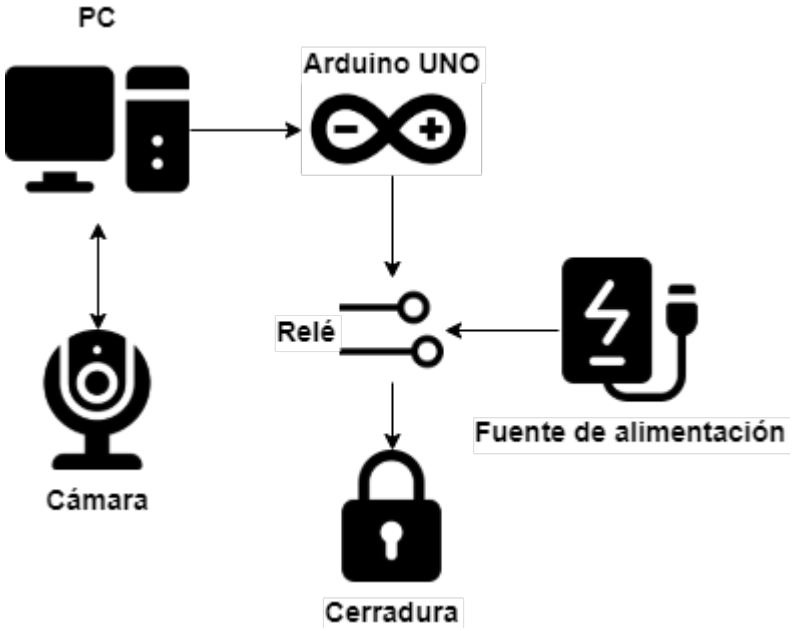


Fig. 1. Diagrama de interacción de los dispositivos de hardware en el sistema identificador facial.

2. La autenticación de la persona, donde se comprueba que no es una fotografía.
3. La apertura de la cerradura utilizando sistemas embebidos.
4. Finalmente el administrador del sistema, donde se gestionan usuarios, roles, etc.

En este trabajo en particular se hizo un enfoque en el proceso de identificación facial utilizando herramientas a partir de imágenes de intensidad (holístico) mediante la utilización de una red neuronal convolucional (CNN). En los resultados se muestran algunas métricas numéricas que evalúan el rendimiento del reconocimiento facial, el cual fue sometido a varias pruebas.

2. Reconocimiento facial

El reconocimiento facial es el proceso de identificar el rostro de una persona relevante mediante un sistema de visión. Ha sido una herramienta crucial de interacción persona-computadora debido a su uso en sistemas de seguridad, control de acceso, videovigilancia, áreas comerciales e incluso se usa en redes sociales como Facebook [4]. El reconocimiento facial parece ofrecer varias ventajas sobre otros métodos biométricos, algunos de los cuales se describen aquí:

Casi todas estas tecnologías requieren alguna acción voluntaria por parte del usuario, es decir, el usuario debe colocar su mano en un escáner para la toma de huellas dactilares o la detección de la geometría de la mano y debe pararse en una posición fija frente a una cámara para la identificación del iris o la retina.

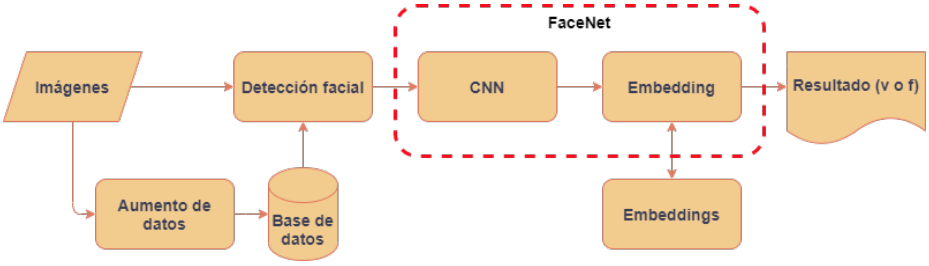


Fig. 2. Diagrama de flujo de información a través del software.

Sin embargo, el reconocimiento facial se puede realizar de forma pasiva sin ninguna acción o participación explícita por parte del usuario, ya que una cámara puede adquirir imágenes faciales a distancia [7].

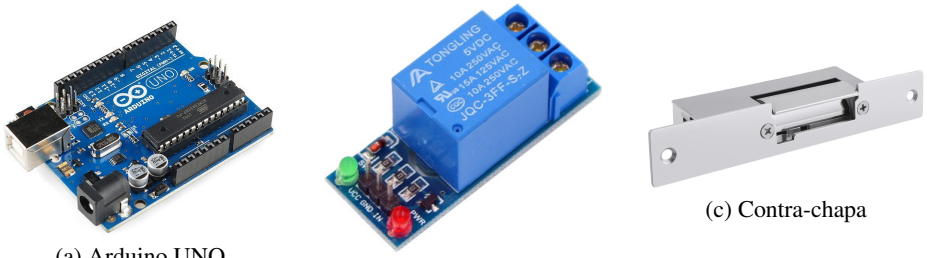
Existen diferentes técnicas para realizar la tarea de reconocimiento facial que se pueden dividir en tres grupos principalmente: 1) métodos que operan con imágenes de intensidad, 2) aquellos que tratan con secuencias de vídeo y 3) aquellos que requieren otros datos sensoriales como información 3D o imágenes infrarrojas [7].

– **Reconocimiento facial a partir imágenes de intensidad**: Los métodos de reconocimiento facial de imágenes de intensidad son dos, el método basado en características y el método holístico:

  • **Basado en características.** En este método, las características locales como ojos, nariz y boca se extraen en primer lugar y sus ubicaciones y estadísticas locales (geométricas y/o de apariencia) se introducen en un clasificador estructural. Un gran desafío para los métodos de este tipo es la restauración"de características, esto es cuando el sistema intenta recuperar características que son invisibles debido a grandes variaciones, p. Ej. Pose de cabeza cuando estamos haciendo coincidir una imagen frontal con una imagen de perfil [12].

  • **Holístico**: A diferencia del método basado en características, este método realiza la tarea de reconocimiento facial utilizando representaciones totales de la imagen, este método se divide en dos grupos diferentes: aproximaciones estadísticas e inteligencia artificial (IA). Las aproximaciones estadísticas realiza el reconocimiento comparaciones de correlación directa entre la cara de entrada y todas las demás caras de la base de datos [7]. Los métodos de IA e basan en la utilización de redes neuronales y aprendizaje maquina para lograr alcanzar altos estándares de calidad en el reconocimiento facial.

– **Reconocimiento facial a partir de secuencias de vídeo**: Un sistema de reconocimiento facial basado en vídeo generalmente consta de tres módulos: uno para detectar el rostro; un segundo para rastrearlo; y un tercero para reconocerlo. La mayoría de estos sistemas eligen unos pocos fotogramas buenos y luego aplican una de las técnicas de reconocimiento de imágenes de intensidad a esos fotogramas para identificar al individuo [7].



(a) Arduino UNO

(b) Relevador

(c) Contra-chapa

Fig. 3. Componentes de hardware utilizados en la cerradura biométrica.

– **Reconocimiento facial a partir de datos sensoriales**: Los métodos de reconocimiento facial a partir de datos sensoriales se empezaron a desarrollar debido a factores de posicionamiento del rostro en la cámara de vídeo, para que los métodos descritos anteriormente funcionen el rostro del usuario debe mostrarse de manera frontal, es decir, en 2D para poder extraer las características del mismo. Con los datos sensoriales es posible obtener imágenes del rostro del usuario en 3D, esto lo podemos dividir en dos técnicas distintas: modelo basado en 3D e infrarrojo. Ambas técnicas utilizan procedimientos bastante parecidos, se diferencian de las tecnologías empleadas para la obtención de la información para procesar.

3. Trabajos relacionados

En el área de cerraduras biométricas, donde características corporales de una persona son la llave de la cerradura, encontramos el trabajo de Baidya et al. [1], 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 desbloquear la puerta o bien emitir señales de indicación auditiva (buzzer) y visual (matriz de leds 4x4).

En los trabajo de Radzi et al. [13] y Balla et al. [2] 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. [17], desarrollaron una sistema de cerradura biométrica basado en reconocimiento facial utilizando el método Viola-Jones, realizando la clasificación utilizando la ecuación de distancia euclidiana para identificar al la persona. A diferencia de los trabajos descritos anteriormente, este sistema esta implementado en una PC convencional y en MatLab, si bien los niveles de seguridad son mucho mas deficientes en comparación con otras opciones, este trabajo presume de practicidad y una buena implementación a lo pretendido por los autores.

En el área de reconocimiento facial se han documentado diferentes trabajos en el tema. Como por ejemplo los trabajos de Parkhi et al. [11], Taigman et al.

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

Fig. 4. Resultados del proceso de aumento de datos.

[15], Cao et al. [3] donde utilizan técnicas de aprendizaje profundo para realizar el reconocimiento facial. En [11] realizan el reconocimiento facial de personas importantes del mundo haciendo uso de servicios web gratis.

Probando diferentes arquitecturas como Fisher Vector Faces, DeepFace, Fusion, DeepID-2,3, FaceNet y FaceNet + Alignment; donde la bases de datos utilizada es YouTube Faces Dataset [16]. En [15] dividen el proceso en tres partes: alineación facial, representación, y la tercera parte corresponde a las métricas de verificación. La alineación facial se encarga de detectar las caras en las imágenes, basándose en puntos de referencia que detectan características correspondientes a la forma de la cara.

En la etapa de representación, se encuentran todos los procesos correspondientes a la extracción de las características que diferencian una cara de otra. Para finalmente, 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. En [3] crea una base de datos de más de 3 millones de imágenes faciales clasificadas en más de 9000 identidades diferentes.

Después, utiliza una línea de tiempo acerca de cómo crear una base de datos con características capaces de clasificar imágenes faciales. Posteriormente se continúa creando una plantilla para el conjunto de prueba enfocada en explorar el rendimiento del reconocimiento de pose y edad. Para finalmente, demostrar que el entrenamiento de las CNN con la nueva base de datos mejora significativamente el rendimiento en comparación con el estado del arte. La base de datos utilizada selecciona nombres de personalidades como artistas, políticos, deportistas, etc. disponibles en Google Image Search.

4. Materiales y métodos

Para explicar el desarrollo del sistema de identificación facial propuesto en este documento primero describiremos 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

Un sistema de cerradura biométrica otorga acceso a los usuarios autorizados mediante la verificación de sus características físicas o de comportamiento únicas, como

Fig. 5. Comportamiento de las imágenes con el proceso Triplet Loss.

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 [10].

El sistema propuesto en este trabajo, utiliza el reconocimiento facial como llave para obtener acceso a un área deseada. En la Figura 1 se observa el diagrama correspondiente de la relación existente entre los componentes físicos que conforman al sistema de cerradura biométrico propuesto.

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. La cámara utilizada en este proyecto esta interna en la PC y es de 1440p (alta definición).

El software del sistema es ejecutado en una PC, para este proyecto se utilizó una computadora con procesador Intel Core i7, 2.10 GHz de velocidad, 6 GB de RAM y sistema operativo Windows 7 Professional de 64 bits. Este software envía ordenes a un dispositivo Arduino UNO (Figura 2a) que se conecta a la PC vía USB para diferir las tareas de bloqueo y desbloqueo de la cerradura electrónica.

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. La cerradura electrónica consiste en una contra-chapa (Figura 2b) que se desbloquea al recibir una señal de 12V, esta señal la recibe desde una fuente de alimentación de 12V, ya que la placa Arduino solo es capaz de enviar voltajes máximos de 5V.

Debido a lo explicado anteriormente, es necesaria la utilización de un relé eléctrico (Figura 2c), este dispositivo hace la función de ”switch”, es decir, 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.

4.2. 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

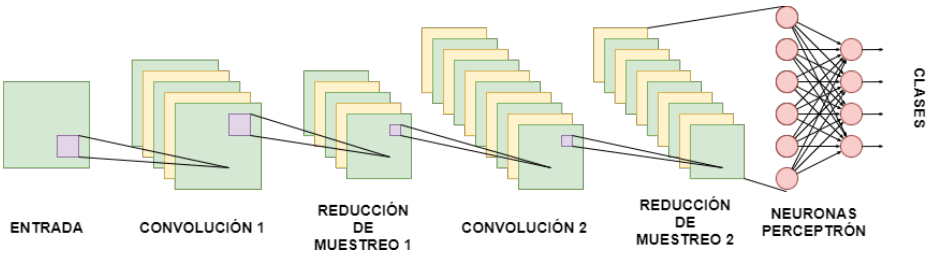


Fig. 6. Arquitectura clásica de una CNN.

desbloqueo o bloqueo de la cerradura biométrica respectivamente. En la Figura 2 se muestra el diagrama de flujo del tratamiento de las imágenes el cual lo podemos dividir en dos bloques.

El primer bloque corresponde a la creación de la base de datos de usuarios que el sistema puede reconocer facialmente. El segundo bloque realiza la tarea de Reconocimiento Facial en tiempo real, este boque se puede subdividir en tres etapas: (i) Detección Facial, (ii) Identificación Facial (usando una CNN) y (iii) Embedding.

4.3. Creación de la base de datos

Para la creación de la base de datos es necesario capturar fotografías frontales de los rostros de los usuarios que se desea tener disponibles en el sistema para su reconocimiento. Previo al cálculo de los Embeddings de estas imágenes, se realiza un aumento de datos (Figura 4).

El aumento de datos consiste en generar dos copias modificadas adicionales a partir de la original (Figura 4a), a la primera se le aplica un ligero cambio de ángulo (Figura 4b) y a la segunda un efecto de acercamiento (Figura 4c). Esto con el fin de fortalecer la base de datos y mejorar el cálculo de las distancias entre las personas.

Como se menciono anteriormente, estas imágenes son procesadas por el detector facial y la FaceNet. El resultado final de este tratamiento a las imágenes disponibles en la base de datos es un archivo llamado Embeddings, que son las distancias euclidianas de todas las identidades en un espacio de 128 dimensiones.

4.4. Reconocimiento facial

Como se menciono anteriormente, este bloque se subdivide en tres etapas. La etapa de detección facial se encarga de detectar y extraer los rostros de una imagen, 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, si una o varias regiones de la imagen corresponden con estos moldes, se extraen estas regiones en forma de coordenadas, si una imagen contiene rostros y son detectados se envían directamente a la siguiente etapa.

Las redes neuronales convolucionales (Figura 6) están compuestas por la capa de extracción de características y la capa de aprendizaje. La capa de extracción de

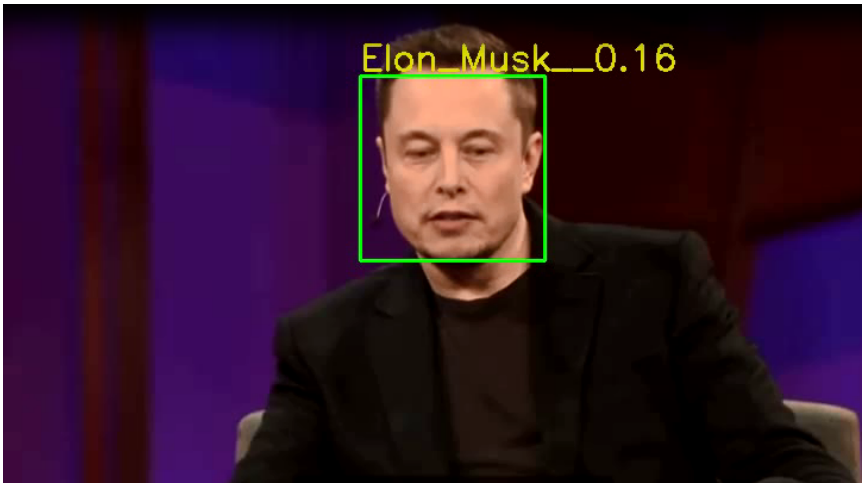


Fig. 7. Ejemplo de un reconocimiento facial exitoso.

características esta formada por capas de convolución y de pooling (reducción) y la capa de aprendizaje regularmente es una red fully conected.

Las CNN estan 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 [16].

Las etapas de Identificación Facial y Embedding se realizan con una herramienta desarrolla por Google llamada FaceNet [14]. 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 CNNs 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 (en su documentación menciona que) logra una efectividad del 99.63 %, siendo una herramienta muy útil y fácil de usar.

El entrenamiento de la FaceNet se realizó utilizando un proceso llamado ”Triplet Loss”(Figura 5). El proceso de Triplet Loss minimiza la distancia entre la base y la muestra para un rostro contenido en la base de datos. Y maximiza la distancia si la muestra no corresponde al rostro a identificar, lo que significa una identidad diferente [9].

Por lo que, una vez generados los embeddings de la base de datos, las imágenes recibidas en tiempo real son procesadas de la misma manera, pero para los fines de reconocimiento, el Embedding generado por la imagen no se almacena, sino que se compara la distancia con cada una de las identidades disponibles en los Embeddings y si la distancia está por debajo de un umbral, se le asigna esa identidad a la imagen (Figura 7).

Tabla 1. Matriz de confusión de los resultados obtenidos.

Real/Predicción	Positivo	Negativo
Positivo	160	0
Negativo	0	40

5. Análisis de resultados

El sistema de reconocimiento facial fue construido, 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 50 personas, de las cuales 40 se encontraban registradas en el sistema para evaluar su correcta identificación y la correcta no identificación de las 10 personas restantes.

Para agilizar las pruebas, se realizó un vídeo con segmentos extraídos de otros vídeos de las 50 personas mencionadas anteriormente, para someterlo al identificador y obtener los resultados. La Tabla 1 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 = 160) los cuales corresponden a la correcta identificación de una persona (asignar la identidad correcta al rostro de una persona). En el cuadrante superior derecho se determinan los resultados falsos-negativos (FN = 0) los cuales corresponden a la incorrecta no identificación de una persona (no asignar una identidad a un rostro disponible en la base de datos).

En el cuadrante inferior izquierdo se determinan los resultados falsos-positivos (FP = 0) los cuales corresponden la incorrecta identificación de una persona (asignar una identidad a un rostro incorrecto). Finalmente, en el cuadrante inferior derecho se determinan los resultados verdaderos-negativos (VN = 40) los cuales corresponden a la correcta no identificación de una persona (no asignar identidad a un rostro no disponible en la base de datos).

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 [16]. La sensibilidad cuantifica qué tan bien el modelo evita los falsos negativos (Ecuación 1). La especificidad cuantifica qué tan bien el modelo evita los falsos positivos (Ecuación 2):

$$\text{Sensibilidad} = \frac{\text{VP}}{\text{VP} + \text{FN}}, \tag{1}$$

$$\text{Especificidad} = \frac{\text{VN}}{\text{VN} + \text{FP}}. \tag{2}$$

Tabla 2. Resultados obtenidos de las mediciones realizadas.

Medida	**Resultado**
Sensibilidad	1.0
Especificidad	0.0
Exactitud	1.0
Precisión	1.0
F1	1.0

La exactitud es el grado de cercanía de las mediciones de una cantidad al valor real de esa cantidad (Ecuación 3) [16]:

$$\text{Exactitud} = \frac{\text{VP} + \text{VN}}{\text{VP} + \text{FP} + \text{FN} + \text{VN}}. \tag{3}$$

El grado en que las mediciones repetidas en las mismas condiciones nos 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 4) [16]:

$$\text{Precision} = \frac{\text{VP}}{\text{VP} + \text{FP}}. \tag{4}$$

En la clasificación binaria, consideramos 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 [16], como se define aquí:

$$\text{F1} = \frac{2\,\text{VP}}{2\,\text{VP} + \text{FP} + \text{FN}}. \tag{5}$$

Vemos puntajes para F1 entre 0.0 y 1.0, donde 0.0 es el peor puntaje y 1.0 es el mejor puntaje que nos 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, vemos que la puntuación F1 se utiliza como una puntuación general sobre el rendimiento de nuestro modelo [16].

En la Tabla 2 se encuentran los resultados alcanzados de las diferentes medidas descritas anteriormente con base en la matriz de confusión obtenida al principio de este capítulo. 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 menciono anteriormente la red neuronal utilizada para este proyecto alcanza un 99.63 % de rostros reconocidos exitosamente, es decir, 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 las pruebas realizadas al sistema, constaron de fragmentos de vídeos que favorecieran al reconocimiento, es decir, aseguramos 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, esto debido a que el sistema supone de un consentimiento por parte del usuario a mostrar su rostro ante la cámara para su identificación.

6. Conclusiones

El diseño de la cerradura biométrica es sencilla y funcional para los fines propuestos en el trabajo. El desempeño del hardware utilizado fue el esperado y se lograron realizar las pruebas correspondientes al sistema de software de una manera excelente, obteniendo resultados extraordinarios a razón de costo/beneficio. Nuestro sistema de reconocimiento facial posee características ideales para su implementación en sistemas de acceso biométricos, ya que brinda altos estándares de fiabilidad, incluso si se utiliza un alto numero de usuarios registrados.

En su conjunto (hardware y software) el sistema propuesto, ofrece una notable opción para quienes requieren de un sistema de cerradura. 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 beneficio de precio se puede ver aumentado significativamente si toman en cuenta que es posible administrar múltiples puertas con un solo dispositivo y solamente colocando una cámara por puerta.

6.1. Trabajo futuro

Existen diferentes áreas de oportunidad derivadas de nuestro sistema que seguramente aumentaran las capacidades de nuestra propuesta.

La primera es migrar a un dispositivo de IoT capaz de soportar sistemas basados en redes neuronales, p. ej., una Jetson Nano de NVIDIA. Esta extensión del proyecto proveería de un aumento de algunas capacidades, p. ej.: implementaciones de IoT, portabilidad, velocidad de respuesta, menor tamaño de prototipo, entre otras.

La segunda es incrustar una etapa de detección de realidad previa a la etapa de reconocimiento facial, esto es debido a que el sistema actual reconoce de igual forma a personas reales y fotografías. Una etapa de detección de realidad, filtraría únicamente rostros de personas reales y desecharía rostros de fotografías impresas o digitales.

Referencias

1. Baidya, J., Saha, T., Moyashir, R., Palit, R.: Design and implementation of a fingerprint based lock system for shared access. In: Proceedings of the IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC). IEEE (2017) doi: 10.1109/ccwc.2017.7868448
2. Balla, P. B., Jadhao, K. T.: IoT based facial recognition security system. In: Proceedings of the International Conference on Smart City and Emerging Technology (2018) doi: 10.1109/icscet.2018.8537344

3. Cao, Q., Shen, L., Xie, W., Parkhi, O.M., Zisserman, A.: VGGFace2: A dataset for recognising faces across pose and age. In: Proceedings of the 13th IEEE international conference on automatic face and gesture recognition (2018) doi: 10.48550/arXiv.1710.08092
4. Coskun, M., Ucar, A., Yildirim, O., Demir, Y.: Face recognition based on convolutional neural network. In: Proceedings of the International Conference on Modern Electrical and Energy Systems (2017) doi: 10.1109/mees.2017.8248937
5. Divya, R. S., Mathew, M.: Survey on various door lock access control mechanisms. In: Proceedings of the International Conference on Circuit, Power and Computing Technologies (2017) doi: 10.1109/iccpct.2017.8074187
6. INEGI. Incidencia delictiva (2021) www.inegi.org.mx/temas/incidencia/
7. Jafri, R., Arabnia, H. R.: A survey of face recognition techniques. Journal of Information Processing Systems, vol. 5, no. 2, pp. 41–68 (2009) doi: 10.3745/JIPS.2009.5.2.041
8. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature, vol. 521, no. 7553, pp. 436–444 (2015) doi: 10.1038/nature14539
9. Ming, Z., Chazalon, J., Luqman, M. M., Visani, M., Burie, J. C.: Simple triplet loss based on intra/inter-class metric learning for face verification. In: IEEE International Conference on Computer Vision Workshops (2017) doi: 10.1109/iccvw.2017.194
10. Nehete, P. R., Chaudhari, J., Pachpande, S. R., Rane, K. P.: Literature survey on door lock security systems. International Journal of Computer Applications, vol. 153, no. 2, pp. 13–18 (2016) doi: 10.5120/ijca2016911971
11. Parkhi, O. M., Vedaldi, A., Zisserman, A.: Deep face recognition. In: Procedings of the British Machine Vision Conference, British Machine Vision Association (2015) doi: 10.5244/c.29.41
12. Parmar, D. N., Mehta, B .B.: Face recognition methods and applications. Journal of Information Processing Systems, vol. 5, no. 2, pp. 41–68 (2009) doi: 10.3745/JIPS.2009.5.2.041
13. Radzi, S. A., Alif, M. K., Athirah, Y. N., Jaafar, A. S., Norihan, A. H., Saleha, M. S.: IoT based facial recognition door access control home security system using raspberry pi. International Journal of Power Electronics and Drive Systems, vol. 11, no. 1, pp. 417 (2020) doi: 10.11591/ijpeds.v11.i1.pp417-424
14. Schroff, F., Kalenichenko, D., Philbin, J.: FaceNet: A unified embedding for face recognition and clustering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015) doi: 10.1109/cvpr.2015.7298682
15. Taigman, Y., Yang, M., Ranzato, M., Wolf, L.: DeepFace: Closing the gap to human-level performance in face verification. IEEE Conference on Computer Vision and Pattern Recognition (2014) doi: 10.1109/cvpr.2014.220
16. Wolf, L., Hassner, T., Maoz, I.: Face recognition in unconstrained videos with matched background similarity. In: Proceedings of the Computer Vision and Pattern Recognition Conference (2011) doi: 10.1109/CVPR.2011.5995566
17. Yedulapuram, S., Arabelli, R., Mahender, K., Sidhardha, C.: Automatic door lock system by face recognition. IOP Conference Series: Materials Science and Engineering, vol. 981, no. 3 (2020) doi: 10.1088/1757-899x/981/3/032036