

Reconocimiento facial basado en aprendizaje profundo para aplicaciones de seguridad automotriz

Deep learning-based face recognition for automotive security applications

José Manuel Jiménez García¹, Juan P. Amézquita Sánchez¹, Martín Valtierra Rodríguez^{1,*}

¹ ENAP -Research Group, CA-Sistema Dinámicos y Control, Universidad Autónoma de Querétaro (UAQ), Campus San Juan del Río, Río Moctezuma 249, Col. San Cayetano, San Juan del Río 76807, México.

*Autor de correspondencia: martin.valtierra@enap-rg.org

Resumen

En el ámbito de la seguridad automotriz, las soluciones varían desde métodos básicos como candados y cerraduras hasta sistemas avanzados que emplean tecnologías de reconocimiento biométrico, como huella digital y detección facial. Este estudio presenta un algoritmo innovador de reconocimiento facial que autentica a los usuarios mediante el análisis de sus rostros en tiempo real. Utilizando el algoritmo Viola-Jones, se capturan y recortan las imágenes faciales, que luego se emplean para entrenar y probar la Red Neuronal Convolutiva AlexNet. Como resultado, el sistema logra reconocer al usuario y permite el encendido del automóvil a través de un relevador. Los resultados obtenidos demuestran una efectividad promedio del 77.68%, alcanzando un máximo del 90% en algunos casos, lo que demuestra el potencial de esta aplicación en el contexto de la seguridad automotriz.

Palabras clave: AlexNet, Reconocimiento Facial, Redes Neuronales Convolutivas, Seguridad automotriz, Viola-Jones.

Introducción

Durante las últimas décadas la industria automotriz ha tenido un enorme avance tecnológico. Año tras año se han encontrado nuevas tecnologías en cuanto a la comodidad dentro del automóvil, tecnologías para un mejor desempeño y de igual manera tecnologías novedosas para la seguridad del automóvil tanto para alguna situación de percance como para la seguridad en la apertura de este. Algunos países de Latinoamérica como México y Brasil se han posicionado en los primeros lugares de robo vehicular, este delito se realiza a través del uso de la fuerza (Editorial, 2018). De acuerdo con la Organización de los Estados Americanos, pese a que el índice de robo a auto particular decreció en el último año, México sigue encabezando el listado de naciones en América Latina por esta infracción, estando en segundo lugar Argentina, seguido de Uruguay, Venezuela y Costa Rica (OEA)(Atracion, 2013). Durante el último año, se reportaron en el país 62 mil 102 vehículos robados, esto durante periodo de abril del 2021 a marzo del 2022, donde fueron robados diariamente 170 unidades (Cruz, 2022).

Anteriormente se han propuesto diferentes trabajos buscando la seguridad dentro del automóvil, los cuales comenzaron desde la creación de dispositivos que inmovilizan una parte del automóvil (bastones, seguros de rueda, cortadores de corriente) y que son encontrados de manera comercial a un muy bajo costo como es el caso del bastón el cual su precio oscila entre los \$200.00 hasta los \$600.00 MXN (AutoZone, 2023). El mismo caso es para los inmovilizadores los cuales son aún más caros que los bastones llegando a costar entre \$1,800.00 a \$4,200.00 MXN (Libre, 2023), sin embargo cuentan con la misma confiabilidad y el proceso para colocarlos es mucho mayor debido al peso y tamaño de estos.

En los últimos años, se han reportado sistemas de seguridad automotriz de bajo costo y utilizando tecnología de reconocimiento facial, comenzado, por ejemplo, con sistemas como el trabajo elaborado por Salazar, (2019), el cual se desarrolló un sistema de seguridad automotriz con autenticación de reconocimiento facial mediante visión artificial. En otra investigación, Quiroz, (2020), realizó un prototipo de sistema para la seguridad del automóvil mediante reconocimiento de rostro y huella digital, para ello utilizó la librería Open CV para la detección de rostro y el microprocesador Raspberry en el cual va montado también un módulo para detección de huella digital. Otra aportación en ésta área de la seguridad es el trabajo de Jiménez, (2020) quien implementó un sistema de reconocimiento de rostros el cual alertaba al usuario en caso de que alguien no autorizado entrara en su vehículo.



Aunque los trabajos descritos anteriormente presentan contribuciones significativas, el explorar y realizar sistemas basados en redes neuronales convolucionales o profundas podría incrementar la certeza y fiabilidad en las aplicaciones relacionadas con reconocimiento facial debido a las altas capacidades que se han reportado en la literatura. Existen algunos trabajos en esta área como el artículo realizado por William et al. (2019), en el cual se propuso e implementó un sistema automático de encendido-apagado del motor de un automóvil basado en el reconocimiento facial utilizando Redes Neuronales Artificiales (ANN), usando procesamiento de las imágenes y el algoritmo Viola-Jones junto con el método Canny para la detección de rostros. Para el entrenamiento utilizaron 10 fotos de dos personas diferentes, el resultado que obtuvieron de la detección de rostros fue el dato de entrada hacia la tarjeta Arduino, el cual activaba un relevador para el encendido del automóvil (William et al., 2019). Bhanse y Jaybhaye (2018), explican cómo desarrollaron un sistema de detección y seguimiento de rostros, donde el propósito del proyecto fue activar un relevador para que éste a su vez activara el dispositivo de apertura de un automóvil. Ellos utilizaron el Algoritmo Viola-Jones junto a la librería AdaBoost para realizar la detección de los rostros, donde el procesamiento de las imágenes se realizó mediante un microprocesador Raspberry PI utilizando programación Python y la librería open CV. Treviño, (2022), desarrolló un sistema para la detección de vehículos y peatones, utilizando redes neuronales convolucionales y algoritmos de búsqueda, logrando la implementación de un algoritmo metaheurístico-híbrido. A partir de los trabajos revisados, es evidente que ya existen diferentes propuestas de seguridad basadas en redes neuronales convolucionales, sin embargo, aún hay algunas oportunidades de investigación, principalmente considerando la exploración de redes pre-entrenadas como es el caso de la Google Net, ResNet o AlexNet, entre otras, las cuales, debido a su amplia estructura, tienen el potencial de una alta eficiencia.

En este trabajo de investigación se implementó una metodología basada en el aprendizaje de la red neuronal convolucional AlexNet para el reconocimiento facial con el fin de obtener un sistema de seguridad mediante el reconocimiento facial para la autorización de la ignición del automóvil. La efectividad del algoritmo propuesto se evalúa mediante la respuesta del sistema a diferentes entradas de rostros. Para la detección y segmentación del rostro se utiliza el algoritmo de Viola-Jones. Todo el sistema está integrado en una interfaz gráfica de usuario desarrollada en Matlab. Esta interfaz integra las capturas de la base de datos, la etapa de entrenamiento y la etapa de prueba. El sistema es probado fuera de línea y en línea dentro de un automóvil.

Marco Teórico

A continuación, se presentan los elementos teóricos precisos para llevar a cabo la metodología presentada, dando una descripción de los diferentes algoritmos empleados para adquirir las imágenes, procesar las imágenes y concluir con el entrenamiento en la red neuronal convolucional AlexNet.

Algoritmo Viola Jones.

Como sugiere el nombre, Paul Viola y Michael Jones introdujeron el algoritmo Viola-Jones Face ID (Face identification o identificador de rostro) en 2001. Este algoritmo fue el primero en proporcionar una tasa de detección de objetos relativamente alta (Rodríguez, 2017). Uno de sus mayores usos es la localización de objetos en tiempo real, pero su mayor uso es el reconocimiento facial. El porcentaje de localización de este algoritmo es comparativamente alta y muy por debajo para obtener inexistentes positivos, generando un algoritmo robusto y con ello procesar imágenes rápidamente. Su principal objetivo es la detección de rostros más no la afirmación, es decir, distinguir los rostros de las no caras. Esta sería la primer etapa del reconocimiento facial (Viola & Jones, 2001).

El algoritmo está basado en el cotejo de áreas rectangulares de la imagen, las llamadas características Haar-Like, las cuales se obtienen a partir de una efigie integral. Estos clasificadores, cuentan con una precisión ligeramente superior a la probabilidad, están conectados en cascada, utilizan el algoritmo de entrenamiento apoyado en AdaBoost con el cual se logra un alto rendimiento en la detección y una alta discriminación en los iniciales períodos (Gonzalez & Velásquez, 2019).



Características Haar-Like

Se describe como Haar-Like a una particularidad simple rectangular que se utiliza como una entrada para el clasificador en cascada. Observando la figura 1, se ejemplifican algunos filtros basados en esta característica. Utilizando cada uno de estos filtros. Por un área específica de la imagen, las adiciones de píxeles debajo de las áreas blancas son restadas de las adiciones de píxeles debajo de las áreas negras. Esto significa que el peso de las áreas blancas y negras se consideran como "1" y "-1", respectivamente (Viola & Jones, 2001).

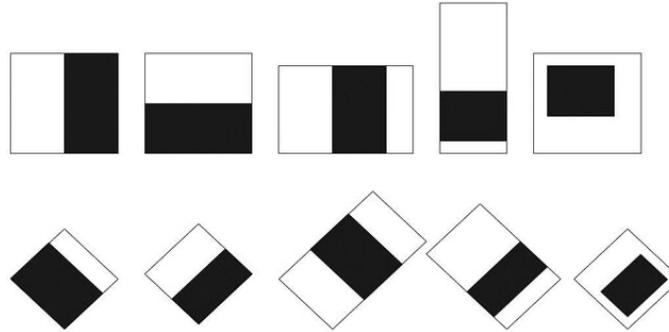


Figura 1. Filtros basados en característica Haar-Like
Fuente: Moghimi et al., 2018.

Dichas particularidades son características muy simples las cuales se indagaron en las representaciones pictográficas y que consisten en variaciones de la luz entre áreas rectangulares contiguas. Las particularidades se denotan por los rectángulos y su lugar referente a la ventana de búsqueda y reciben un valor numérico a través del cotejo que evalúan (Moghimi et al., 2018). En el artículo de Viola-Jones, las particularidades se definen sobre una ventana de exploración elemental de 24x24 píxeles, lo cual genera más de 180000 particularidades posibles (Gonzalez & Velásquez, 2019).

Imagen Integral

Se define como un método para el cálculo de la particularidad Haar-Like. Ese método consiste en sumar cada uno los valores de píxeles en el primer cuadrante superior de izquierda a derecha en la posición (x, y), tal y como se puede apreciar en la figura 2. El área de color negro que se ve en la figura 2 muestra la imagen integral.

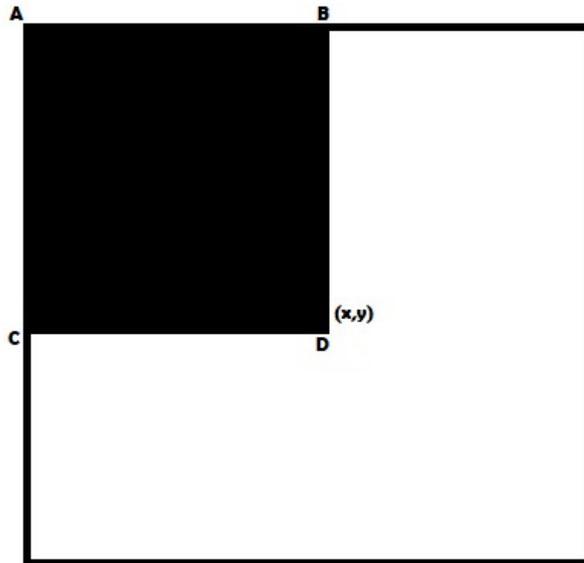


Figura 2. Ejemplo de imagen integral.
 Fuente: Elaboración propia.

La localización (x,y) , es la suma de los píxeles de la parte superior izquierda, para calcularlos se utiliza la ecuación 1 mostrada a continuación:

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x',y'), \quad [1]$$

Donde $ii(x,y)$ es la imagen integral con coordenadas (x,y) y $i(x',y')$ es el punto original. Usar la imagen integral para la suma de cualquier área rectangular es extremadamente eficiente, como se muestra en la figura 2. La suma de píxeles en la región del rectángulo de color negro es posible de deducir siguiendo la ecuación 2 (Zhang & Zhang, 2010):

$$\sum_{(x,y) \in ABCD} i(x,y) = ii(D) + ii(A) + ii(B) + ii(C) \quad [2]$$

la cual solo requiere cuatro referencias de matriz, es decir las posiciones A,B,C,D que se muestran en la figura 2.

Algoritmo AdaBoost

Se conoce como “boosting” o “impulso” al método para encontrar una hipótesis al combinar muchas de estas hipótesis “débiles”, cada una con una puntualidad templada. El método AdaBoost (Adaptative Boosting) generalmente se observa como el primer paso hacia algoritmos de boosting más prácticos (Zhang & Zhang, 2010). El algoritmo Viola-Jones utiliza el AdaBoost como una mezcla de una serie de clasificadores generando con esto una cadena de filtros. Cada uno de estos filtros contienen una serie de clasificadores “débiles” (Moghimi et al., 2018). Si cada uno de los filtros arroja una falla, esta área es tomada como un no rostro, en cambio, si el filtro detecta un área de la imagen como un rostro entonces se procede a evaluar con el siguiente filtro en la cadena (Gonzalez & Velásquez, 2019). Si la imagen pasa con éxito por cada uno de los filtros entonces es catalogada como un rostro.

La clasificación final estará dada por la combinación lineal de la clasificación inicial débil, la cual se rige por las siguientes ecuaciones y reglas (Viola & Jones, 2001):

- Se tiene una imagen $(x_1, y_1) \dots (x_n, y_n)$ donde $y_i = 0,1$ para un resultado negativo o positivo respectivamente.



- Inicializar los pesos $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ para $y_i = 0,1$ respectivamente, donde m y l son el número de negativos y positivos obtenidos.
- Para $t=1, \dots, T$:
 1. Normalizar los pesos

$$W_{t,i} \leftarrow \frac{W_{t,i}}{\sum_{j=1}^n W_{t,j}} \quad [3]$$

Donde w_t es la distribución de la probabilidad.

En cada particularidad j , se genera un clasificador h_j el cual está restringido por una simple particularidad. El error se evalúa respecto a w_t .

$$\epsilon_j = \sum_i w_i |h_j(x_i) - y_i| \quad [4]$$

2. Se escoge el clasificador con menor índice de error.
3. Se renuevan los pesos.

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$$

Donde $e_j = 0$ si el ejemplo x_i se cataloga comedidamente y 1 en caso opuesto.

- El clasificador final es:

$$h(x) = \begin{cases} 1 & \text{si } \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & \text{en caso contrario} \end{cases} \quad [5]$$

donde $\alpha_t = \frac{1}{\beta t}$

Clasificador en cascada

El clasificador en cascada se basa en el principio de que en vez de que sea un único clasificador mediante el algoritmo AdaBoost, se construyen varios clasificadores más pequeños y más eficaces lo cual ocasiona que muchos resultados negativos se rehagan, es decir las ventanas donde no se detecte un rostro se vuelven a construir, con esto se mantienen casi todas las ventanas positivas (Viola & Jones, 2001). Estos clasificadores más simples son utilizados para refutar la mayoría de las ventanas de búsqueda y considerar aquellas en donde hay mayor posibilidad de encontrar caras. Con esto se obtiene una cascada de clasificadores (figura 3) el cual es entrenado con AdaBoost y posteriormente los valores umbrales son ajustados para mermar los falsos negativos.

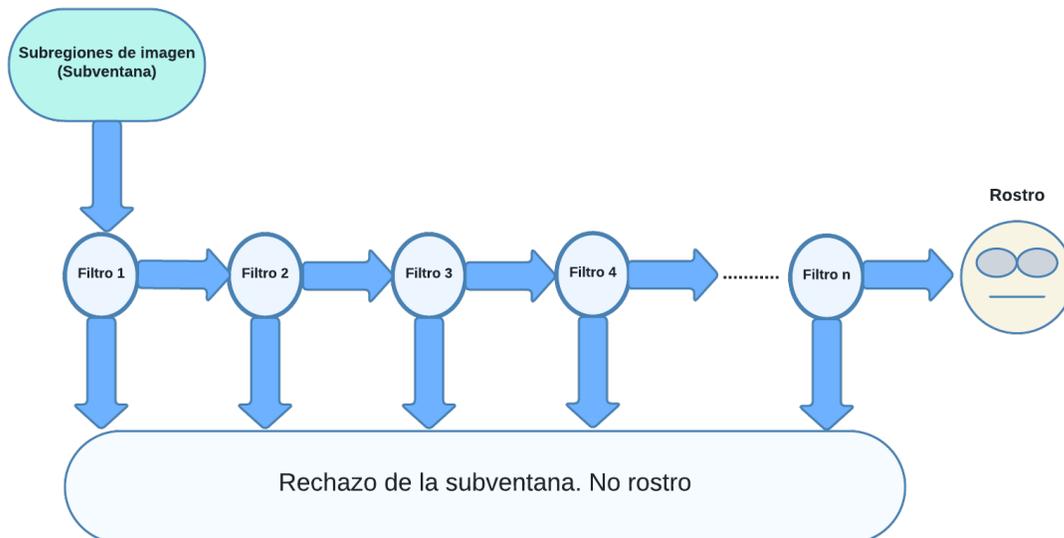


Figura 3. Diagrama de un clasificador en cascada.
 Fuente: Elaboración propia.



El algoritmo elaborado por Viola-Jones tiene 38 etapas y más de 6000 tipologías, en promedio exclusivamente 10 características son evaluadas por ventana de búsqueda (Villanueva & Ramírez Zavala, 2010).

Detección de rostros mediante redes neuronales

El modelo revisa el rostro ordenado mediante zonas de utilidad que exhiben propiedades del color de piel humana. Las propiedades se descubren mediante la base de píxel por píxel. El modelo sugerido se utiliza como un paradigma para procedimientos de reconocimiento facial, sistemas para videovigilancia y métodos para el control de acceso (Rodríguez, 2017).

Redes Neuronales Convolucionales

Diseñar redes neuronales convolucionales (CNN) profundas no es sencillo, debido a que no existe una plantilla con la cual determinar el número de capas, sus tipos o el número de lazos entre ellas. Asimismo, no existe un proceso que simplifique su diseño. Existen aproximaciones matemáticas las cuales aplican para cada una de las redes neuronales convolucionales existentes comenzando por el cálculo de la capa convolucional la cual es denotada por la ecuación 6 (Wang et al., 2022)

$$a_j^l = f \left(b_j^l + \sum_{i \in m_j^l} a_i^{l-1} * k_{ij}^l \right) \quad [6]$$

Donde:

a_j^l = Valor de salida de la j-ésima celda de la capa de convolución l.

b_j^l = Es el sesgo.

k_j^l = Es el kernel de convolución.

Los parámetros se actualizan cuando las actualizaciones de retroalimentación se realizan después de cada ronda de entrenamiento del modelo. Luego se calcula la capa de agrupación la cual es obtenida utilizando la ecuación 7:

$$a_j^l = f \left(b_j^l + \beta_j^l \text{down}(a_j^l, M^l) \right) \quad [7]$$

Donde

b_j^l = Sesgo.

β_j^l = Residual Multiplicativo.

M^l = Es el tamaño de ventana de agrupación.

Down = Denota la relación de agrupación.

AlexNet

La CNN Alex Net (ver figura 4) fue desarrollada por Alex Krizhevsky, Ilya Sutskever y Geoffrey Hinton el año 2012 (Lucero, 2022). Esta CNN tiene una gran importancia y aporte en el área de aprendizaje automático y aplicación en particular como lo es el aprendizaje profundo en el campo de la visión artificial. Cuenta con una estructura semejante Le Net, pero con la diferencia de ser una CNN más profunda, con mayor número de filtros por capa y con capas convolucionales apiladas (Moreno, 2019). Para tener una mayor velocidad, se



usa una función ReLU, que produce un aumento significativo, conservando la precisión original, tiene un mejor rendimiento de entrenamiento Tanh y Sigmoid, usa el Dropout para post-entrenamiento, duplicando el tiempo de entrenamiento con una caída de 0.5.

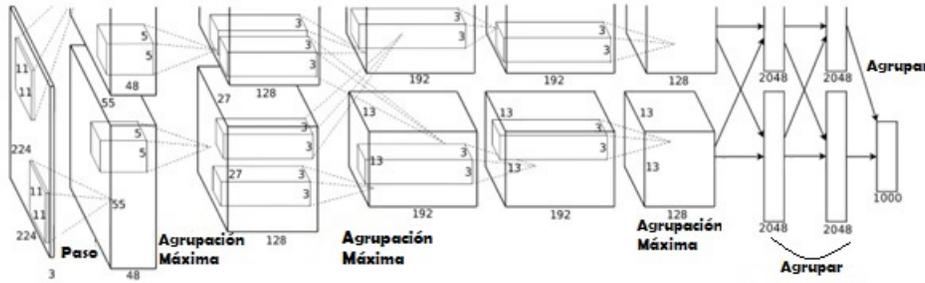


Figura 4. Arquitectura red neuronal convolucional Alex Net
 Fuente: Delbracio & Carbajal, 2017.

Función ReLU

Una de las funciones matemáticas de activación utilizada con gran frecuencia dentro del campo de las redes neuronales convolucionales es la función de unidad lineal rectificadora (Rectified Linear Unit "ReLU" por sus siglas en inglés) (Mellado & Saavedra, 2016), la cual se presenta a continuación en la ecuación 8:

$$f(x) = \max(0, x) \quad [8]$$

siendo x la entrada de la neurona.

Esta función de activación es la más aceptada como un modelo cercano a la biología humana, en comparación con la función sigmoidea la cual es utilizada regularmente en redes neuronales (Glorot et al., 2011).

Función Dropout

El procedimiento Dropout reside en interrumpir la conexión de cierta cantidad de neuronas para cada una de las iteraciones del adiestramiento. Este término se refiere a sacar de una red neuronal, unidades que se encuentran ocultas o también visibles. Esta técnica se utiliza con el fin de evitar un sobreajuste, el cual se presenta cuando tenemos un conjunto de datos reducido para el entrenamiento de la red neuronal convolucional (Chanampe et al., 2019).

En la tabla 1 se muestran las diferentes etapas de entrenamiento para las redes neuronales convolucionales y sus particularidades.

Tabla 1. Algoritmos de ajuste para redes neuronales convolucionales

| Métodos | Particularidades |
|-------------------|---|
| Pre-Entrenamiento | Las redes solo se usan como extractor de características. |
| Afinado | Encargado de ajustar los pesos de una red. |
| Aumento de datos | Este método de aprendizaje amplía artificialmente el conjunto de datos de entrenamiento en imágenes. |
| DropOut | Algoritmo encargado de comprimir el overfitting, separando unidades perceptibles y disimulas de una red neuronal. |

Fuente: Elaboración propia.



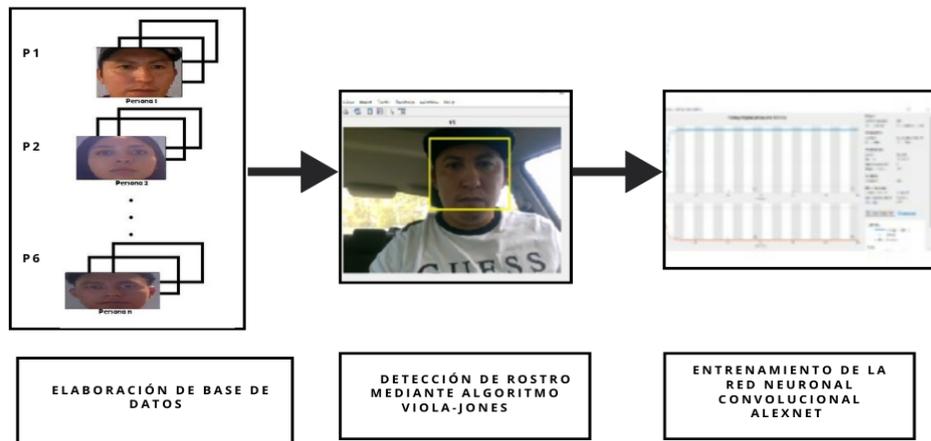
Metodología Propuesta

La figura 2 dividida en sección a y b muestra mediante un diagrama la metodología utilizada para la elaboración del presente proyecto. En este caso es de suma importancia generar una base de datos adecuada al igual que un eficiente entrenamiento de la red neuronal AlexNet para el reconocimiento facial.

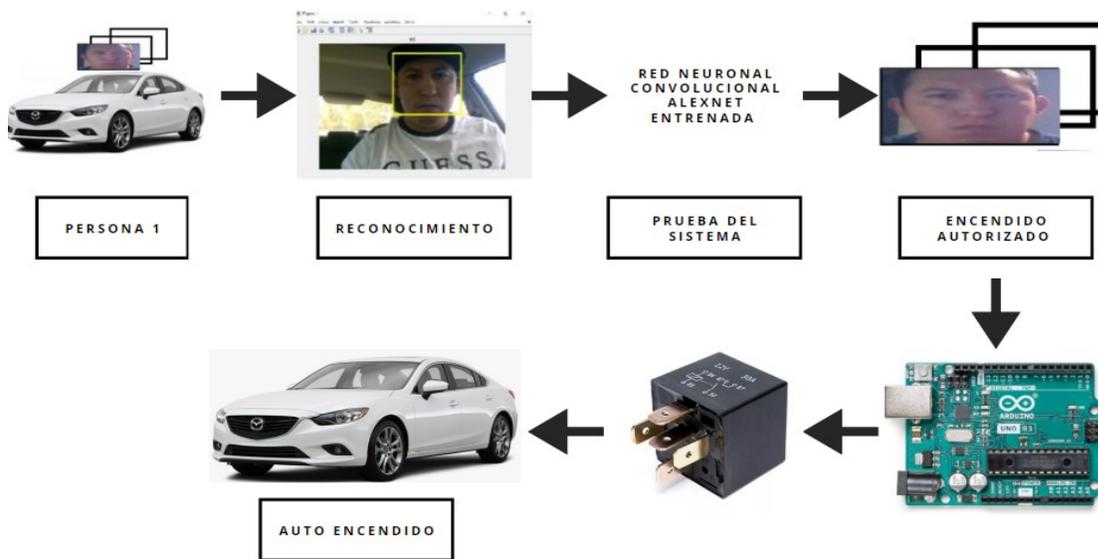
La metodología se divide en 2 principales etapas:

1. La etapa a) o etapa de entrenamiento mostrada en la figura 5. a), consta de tres principales bloques, la primera, mediante el uso del software Matlab, se procede a crear una base de datos. Para el presente proyecto se utilizaron seis personas, las cuales fueron cuatro hombres y dos mujeres, para cada uno de los casos se tomaron 60 muestras, una vez obtenidas las imágenes necesarias, estas se procesan mediante el algoritmo Viola-Jones para encuadrar la parte del rostro de interés para obtener los patrones a detectar. A continuación, con la base de datos obtenida, se realiza el entrenamiento de la red neuronal convolucional AlexNet.
2. La etapa b) o etapa de validación (figura 5. b), se divide en cuatro etapas, la cual consiste en adquirir mediante un video en tiempo real el rostro de la persona dentro del automóvil para poder enviar la señal de activación al sistema de encendido del automóvil, las imágenes de la persona son captadas mediante la cámara del celular y procesadas con el algoritmo Viola-Jones para encuadrar el rostro. Una vez detectado el rostro se procede a comparar con la red neuronal convolucional AlexNet la cual previamente ha sido entrenada, si el rostro detectado es el autorizado para encender el automóvil mediante el uso de la tarjeta Arduino se envía la señal a un relevador para que éste cierre el circuito y permita el flujo de corriente para dar marcha al automóvil.





a)



b)

Figura 5. a) metodología fase de entrenamiento, b) metodología fase de validación
 Fuente: Elaboración propia.

Puesta de Experimento

Para comprobar el funcionamiento de la metodología propuesta para el reconocimiento facial se presenta la siguiente puesta experimental en la figura 6. En forma general se colocó un smartphone en una zona donde se pueda tomar claramente la foto dentro del vehículo y enlazada de forma inalámbrica a la PC (computadora personal) para el envío en tiempo real de la imagen adquirida. Una vez que se concluyeron estos preparativos se procedió a la adquisición de 60 imágenes de cada sujeto de prueba con el fin de fabricar una base de datos; posteriormente, se procede a entrenar la red neuronal convolucional AlexNet para el reconocimiento facial de rostros.

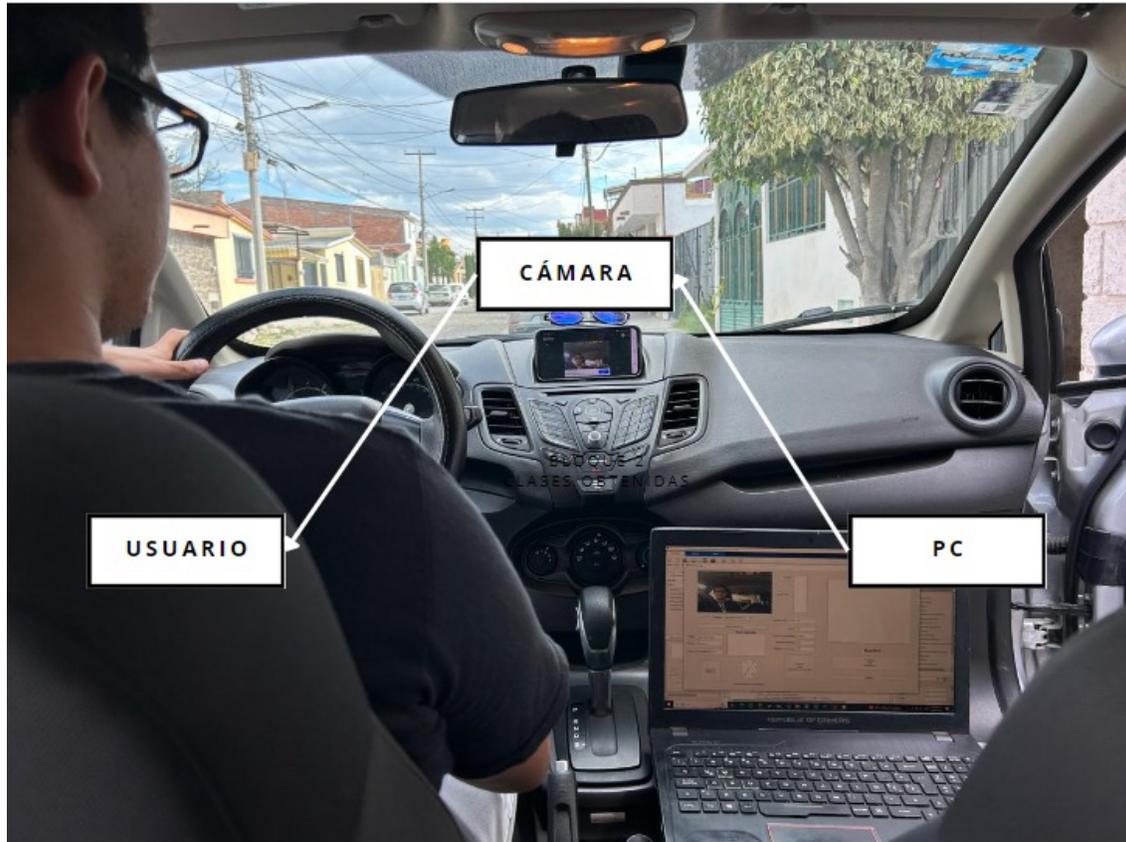


Figura 6. Ilustración del sistema completo, se observa el smartphone colocado en un punto donde el rostro se puede detectar de forma clara con el cual la PC adquiere las fotografías
Fuente: Elaboración propia.

Experimentación y resultados

Para la etapa de entrenamiento se comenzó con la elaboración de una interfaz gráfica en Matlab para con esto adquirir las imágenes que conformarán la base de datos. En la figura 7, se muestra la interfaz gráfica utilizada.

La interfaz está compuesta de cuatro bloques:

- El bloque 1 o bloque de adquisición de imágenes consta de un recuadro en el cual se puede observar en tiempo real las imágenes que se capturan a través de la webcam. En la parte inferior izquierda se encuentra tres cuadros de texto en los cuales se pone el nombre de la carpeta para la base de datos, seguido del nombre del usuario del cual serán adquiridas las imágenes y como último la cantidad de fotos que se tomarán. A un lado se encuentra otro recuadro en el cual aparecerá el rostro detectado por el algoritmo Viola-Jones al igual que un contador para observar cuando el sistema terminó de adquirir las fotografías
- En el segundo bloque o bloque de clases aparece el nombre de la persona fotografiada.
- En el bloque 3 o bloque de entrenamiento de red podemos encontrar cuatro recuadros de texto, en el primero se menciona el nombre de la red neuronal convolucional que se está usando, en este caso es la red neuronal convolucional AlexNet, los siguientes tres recuadros son parámetros para el entrenamiento de la red comenzando por el número de clases con el que se cuenta, seguido de la razón de aprendizaje, el número de épocas y como último el tamaño de paquetes.

- En el último bloque se observa un recuadro en el cual aparecerán las imágenes en tiempo real que detecta la cámara web y en ella se observará la detección del rostro. El nombre del sujeto que detecte aparece por debajo del recuadro.

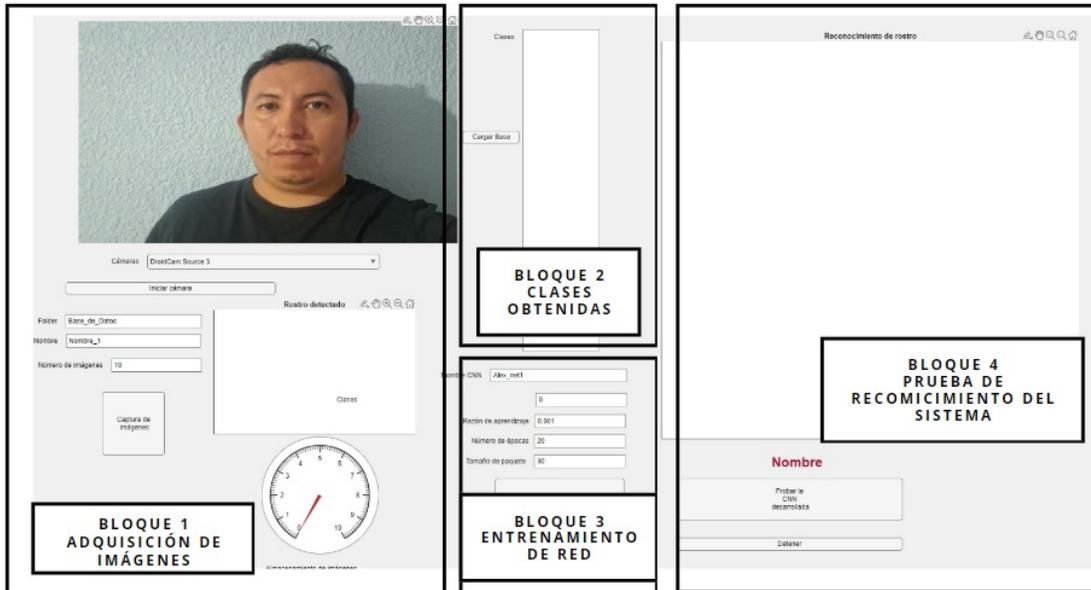


Figura 7. Interfaz gráfica en Matlab para la elaboración de una base de datos
Fuente: Elaboración propia.

Para el presente trabajo de investigación se utilizaron seis personas, las cuales fueron cuatro hombres y dos mujeres, para cada una de las personas se tomaron 60 fotos. En la figura 8 se muestra un ejemplo de cómo se organizaron las fotografías. Una vez generada la base de datos con las cuales se procedió a realizar el entrenamiento de la red neuronal convolucional AlexNet esta se muestra en la figura 9.



Figura 8. Ejemplo de base de datos de la persona 1
Fuente: Elaboración propia.

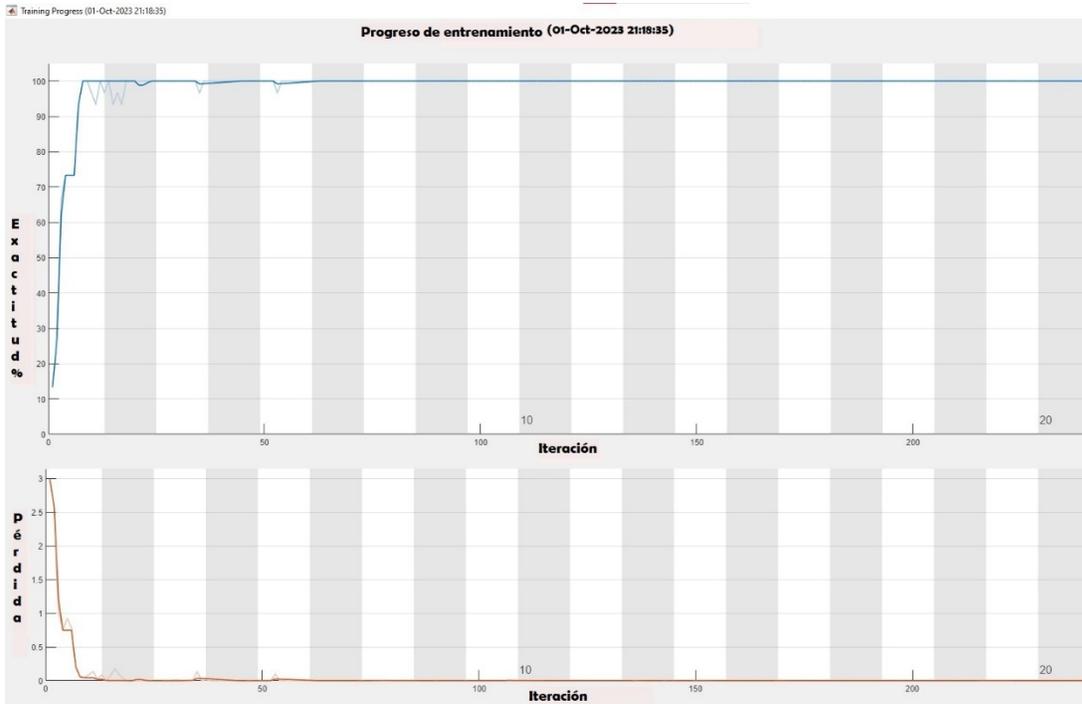


Figura 9. Ilustración del entrenamiento de la red neuronal AlexNet
 Fuente: Elaboración propia.

Una vez realizada la etapa de entrenamiento se procedió a realizar una etapa de validación la cual consiste en que una vez entrenada la red neuronal se toma un video en tiempo real en el cual el sistema detecta el rostro que se le presente, para el caso mostrado en la figura 9. En la tabla 2 se muestran los valores obtenidos, se hicieron diferentes pruebas comenzando con 15 imágenes, luego 25, terminando con 60 imágenes. Para cada prueba se hicieron 60 verificaciones, después se sacó el porcentaje de acierto de cada una de las pruebas y al final se realizó el porcentaje promedio de todos los experimentos realizados.

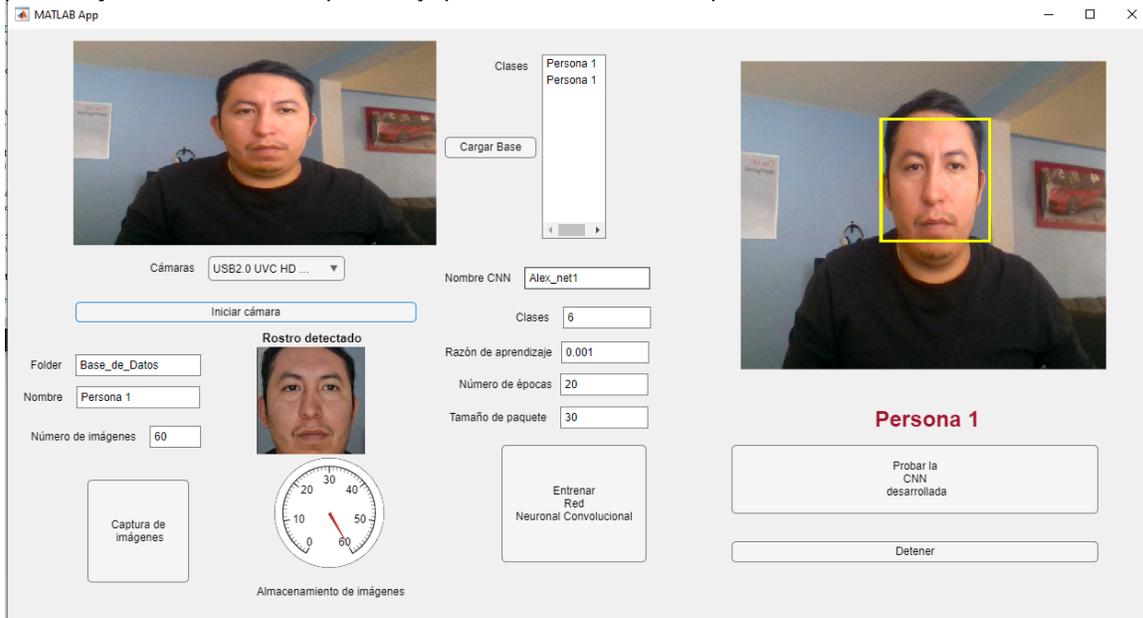


Figura 10. Validación del rostro
 Fuente: Elaboración propia.

Tabla 2. Tabla de experimentación.

| Clase | No. de Imágenes | Porcentaje de acierto | | No. de imágenes | Porcentaje de acierto | | No. de imágenes | Porcentaje de acierto | |
|-----------|-----------------|-----------------------|-------|-----------------|-----------------------|-------|-----------------|-----------------------|-------|
| Persona 1 | 15 | 63.33 | 38/60 | 25 | 70 | 42/60 | 60 | 86.6667 | 54/60 |
| Persona 2 | 15 | 73.33 | 44/60 | 25 | 80 | 48/60 | 60 | 93.33 | 56/60 |
| Persona 3 | 15 | 75 | 45/60 | 25 | 80 | 48/60 | 60 | 93.33 | 56/60 |
| Persona 4 | 15 | 60 | 36/60 | 25 | 75 | 45/60 | 60 | 86.6667 | 52/60 |
| Persona 5 | 15 | 66.6667 | 40/60 | 25 | 76.6667 | 46/60 | 60 | 83.33 | 50/60 |
| Persona 6 | 15 | 70 | 42/60 | 25 | 78.33 | 47/60 | 60 | 86.6667 | 54/60 |
| Promedio | | 68.05445 | | | 76.6661167 | | | 88.3316833 | |

Fuente: Elaboración propia

Conclusiones

En la presente investigación se presentó una metodología basada en el algoritmo de Viola-Jones para la detección y segmentación del rostro, y en el aprendizaje profundo de una red neuronal convolucional AlexNet para el reconocimiento facial de una persona aplicado al sector de seguridad automotriz. Con la finalidad de demostrar la eficiencia del algoritmo empleado a la hora de reconocer el rostro de una persona, se generó y utilizó una base de datos con rostros de 6 personas, las cuales podrían ser los posibles conductores autorizados del automóvil. En la experimentación se obtuvieron diferentes resultados, incluyendo valores arriba del 90%, pero promediando un porcentaje final de acierto del 77.68% a la hora de realizar las pruebas de reconocimiento facial. Es importante mencionar que en algunos casos se encontraron confusiones con patrones similares entre sujetos de prueba debido a que ellos tenían rasgos similares por su parentesco. Esto generó un resultado satisfactorio, pero con varias prospectivas para su mejora como el uso de mayores filtros y parámetros para aumentar el porcentaje de asertividad y bajar el porcentaje de error en el sistema. De igual manera es necesario añadir al sistema un filtro para detectar profundidad, esto ayuda al sistema a ser más seguro debido a que ya no se encontrará vulnerable a ser eludido mediante una fotografía del usuario autorizado para el encendido. También se buscará ampliar el estudio agregando perturbaciones como lo son lentes y gorras.

Para el entrenamiento de la red en trabajos futuros se propone el uso de la función de aumento de datos, la cual aumenta el número de muestras de la base de datos dando diferentes características, como lo es la rotación de las imágenes en rangos de 180 grados, al igual que rotación hacia los lados en el rostro y con ello entrenar de una mejor manera a la red, adicionalmente se plantea la integración del sistema de seguridad físicamente en el automóvil al igual que el relevador para dar el corte de la corriente en el encendido.

Agradecimientos

Se reconoce y agradece al "Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCYT) – México", por financiar parcialmente esta investigación a través de la beca 1242656 otorgada a José Manuel Jiménez - García, y las becas 296574 y 253652 otorgadas a Martín Valtierra-Rodríguez y Juan P. Amézquita-Sánchez, respectivamente, a través del "Sistema Nacional de Investigadoras e Investigadores (SNII)–CONAHCYT–México" Adicionalmente, se reconoce y agradece el apoyo otorgado por parte del Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) a través del Proyecto No. 34/2018 del Programa "Investigadoras e Investigadores por México" del CONAHCYT (Cátedras CONAHCYT).



Referencias

- Atracion. (2013). *En ningún país de latinoamérica se roban mas autos que en México*. <https://www.atraccion360.com/en-ningun-pais-de-latinoamerica-se-roban-mas-autos-que-en-mexico>
- AutoZone. (2023). *AutoZone*. https://www.autozone.com.mx/alarma-y-seguridad/candado-de-seguridad/p/ningbo-motor-candado-rosa-de-2-ganchos-para-volante/1006969_0_0
- Bhanse, V. K., & Jaybhave, M. D. (2018). Face Detection and Tracking Using Image Processing on Raspberry Pi. *Proceedings of the International Conference on Inventive Research in Computing Applications, ICIRCA 2018, Icirca*, 1099–1103. <https://doi.org/10.1109/ICIRCA.2018.8597246>
- Chanampe, H., Aciar, S., De La Vega, M., Luis, J., Sotomayor, M., Carrascosa, G., Lorefice, A., Nacional, U., & Juan, S. (2019). Modelo de redes neuronales convolucionales profundas para la clasificación de lesiones en ecografías mamarias. *XXI Workshop de Investigadores En Ciencias de La Computación, Cidi*, 971–974. <http://sedici.unlp.edu.ar/handle/10915/77381>
- Cruz Ariadna. (2022). *Dónde se robaron más autos este año en México*. <https://www.eluniversal.com.mx/autopistas/donde-se-robaron-mas-autos-este-ano-en-mexico>
- Delbracio, M., Lezama, J., & Carbajal, G. (2017). *APRENDIZAJE PROFUNDO PARA VISIÓN ARTIFICIAL*. https://www.fing.edu.uy/sites/default/files/cursos/2019/anexos/35006/aprendizaje_profundo_para_vision_artificial_a..pdf
- Editorial. (2018). *México y Brasil lideran el robo de autos*. <http://somosseguros.com/cuales-los-dos-paises-donde-mas-autos-se-roban/#:~:text=México y Brasil lideran el,Venezuela%2C Uruguay y Costa Rica.>
- Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. *Journal of Machine Learning Research*, 15(January), 315–323. <https://www.researchgate.net/publication/215616967>
- Gonzalez, H., & Velásquez, S. (2019). Reconocimiento Facial Utilizando Viola-Jones Y Patrones Binarios Facial Recognition Using Viola-Jones and Binary Patterns. *Centro de Investigación de Redes Neuronales Artificiales y Robotica (CIRNAR)*, 23, 57–63. https://www.academia.edu/71107099/Facial_Recognition_Using_Viola_Jones_and_Binary_Patterns
- Jiménez, M. (2020). *Reconocimiento facial como medida de seguridad para alertar el robo de automoviles* [Universidad Autónoma del Estado de Mexico]. <http://hdl.handle.net/20.500.11799/109601>
- Libre, M. (2023). *Mercado Libre*. <https://listado.mercadolibre.com.mx/inmovilizadores-para-ruedas-de-autos>
- Lucero, S. M. (2022). *Comparación entre niños obesos y con normopeso vía análisis del hipotálamo con redes neuronales*. [BENEMERITA UNIVERSIDAD AUTONOMA DE PUEBLA]. <https://hdl.handle.net/20.500.12371/16510>
- Mellado, D., & Saavedra, C. (2016). Aplicación de Deep Learning para la detección de Apnea e Hipoapnea del Sueño basado en Señales de Polisomnografía a Aplicación de Deep Learning para la detección de Apnea e Hipoapnea ~ basado en Señales del Sueño de Polisomnografía. *IX Congreso Anual de Ingeniería Biomédica, October*. <https://www.researchgate.net/publication/327446205%0AAplicación>
- Moghimi, M. M., Nayeri, M., Pourahmadi, M., & Moghimi, M. K. (2018). Moving Vehicle Detection Using AdaBoost and Haar-Like Feature in Surveillance Videos. *“International Journal of Imaging,”* 18(October 2019). <http://arxiv.org/abs/1801.01698>
- Moreno, Á. (2019). Clasificación de Imágenes usando redes neuronales convolucionales en Python [Escuela Técnica Superior de Ingeniería Universidad de Sevilla]. In *Universidad de Sevilla*. <https://idus.us.es/bitstream/handle/11441/89506/TFG-2402-ARTOLA.pdf?sequence=1&isAllowed=y>
- Quiroz, D. E. (2020). *Prototipo de sistema antirrobo de automóviles basado en reconocimiento facial y huellas dactilares*. [Andrés Bello]. <http://repositorio.unab.cl/xmlui/handle/ria/18238>
- Rodríguez, K. E. (2017). *Análisis e Implementación del algoritmo de detección facial de Viola-Jones* [UNIVERSIDAD TECNICA DEL NORTE]. <http://repositorio.utn.edu.ec/handle/123456789/7315>
- Salazar Christian. (2019). *Diseño e Implementación de un sistema de seguridad para un automóvil con autenticación por reconocimiento facial utilizando técnicas de visión artificial* [ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO.]. <https://doi.org/http://dspace.espace.edu.ec/handle/123456789/5008>
- Treviño, G. (2022). *Algoritmo para la detección de vehículos y peatones combinando CNN´s y técnicas de búsqueda*. [Universidad Autónoma de Queretaro]. <https://ri-ng.uaq.mx/handle/123456789/3685>



- Villanueva, M. G., & Ramírez Zavala, S. (2010). Integración de OpenCv en una Raspberry Pi: Sistema de Detección de Rostro vía WEB. *Facultad de Ingeniería Eléctrica de La Universidad Michoacana de San Nicolás de Hidalgo, July*. <https://www.researchgate.net/publication/252627644%0AIntegración>
- Viola, P., & Jones, M. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. *IEEE Journal of Biomedical and Health Informatics*, 7. <https://doi.org/10.1109/CVPR.2001.990517>
- Wang, Y., Wang, J., Zhang, S., Xu, D., & Ge, J. (2022). Remaining Useful Life Prediction Model for Rolling Bearings Based on MFPE–MACNN. *Entropy*, 33. <https://doi.org/10.3390>
- William, L., Winda, A., Satrio, D., Sofyan, T., & Solihin, M. I. (2019). Automotive start-stop engine based on face recognition system. *E3S Web of Conferences*, 130, 1–15. <https://doi.org/10.1051/e3sconf/201913001020>
- Zhang, C., & Zhang, Z. (2010). A Survey of Recent Advances in Face Detection. In *Learning* (Issue June). <https://doi.org/10.1.1.167.5270>

