

Reconocimiento Facial y Detección de Somnolencia en Conductores

*Drivers' Facial Recognition and Drowsiness
Detection*

Alba Cruz Torres

D^a. **Pino Caballero Gil**, con N.I.F. 45.534.310-Z Catedrática de Universidad adscrita al Departamento de Ingeniería Informática y de Sistemas de la Universidad de La Laguna, como tutora

D^a. **Candelaria Hernández Goya**, con N.I.F. 45.441.714-Q profesora Titular de Universidad adscrita al Departamento de Ingeniería Informática y de Sistemas de la Universidad de La Laguna, como cotutora

C E R T I F I C A (N)

Que la presente memoria titulada:

“Reconocimiento Facial y Detección de Somnolencia en Conductores”

ha sido realizada bajo su dirección por D^a. **Alba Cruz Torres**, con N.I.F. 79.073.834-H.

Y para que así conste, en cumplimiento de la legislación vigente y a los efectos oportunos firman la presente en La Laguna a 07 de septiembre de 2022.

Agradecimientos

Agradezco a mi familia y mis amigos por todo el apoyo que me han brindado a lo largo de este último año, ya que sin ellos no habría sido posible la finalización de este proyecto.

También agradecer a mi pareja, quien ha estado conmigo en todos los momentos difíciles y me ha apoyado en todas mis decisiones.

Asimismo, quisiera dar las gracias a mi padre, pues sé que siempre me acompañará en mis éxitos y en mis derrotas.

Finalmente, me gustaría agradecer a la empresa Nokia Spain SA., por permitirme participar en el desarrollo del proyecto IMMINENCE.

Licencia



© Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional.

Resumen

Los accidentes de tráfico causan continuos fallecimientos en todo el mundo. A medida que se va desarrollando nueva tecnología aplicable y se van endureciendo las sanciones, se consigue reducir progresivamente el número de víctimas mortales, pero aun así hoy en día sigue habiendo demasiados fallecidos en las carreteras. Por ese motivo, en este trabajo se ha realizado una investigación de algunas de las soluciones existentes en el mercado para aumentar la seguridad vial en las carreteras mediante el reconocimiento facial de los conductores y la detección de signos de somnolencia. A partir de dicho análisis se ha estudiado el diseño óptimo para cumplir el objetivo establecido. Este documento presenta un estudio de técnicas de Machine Learning y herramientas de detección de rostros aplicadas a un programa para el reconocimiento facial y la detección de somnolencia en conductores.

Palabras clave: seguridad vial, reconocimiento facial, identificación de somnolencia, biometría

Abstract

Traffic accidents cause constant deaths all over the world. As new applicable technology is developed, and sanctions are tightening, the number of fatalities is progressively reduced, but even so there are still too many deaths on the roads. For this reason, in this work an investigation of some of the existing solutions in the market has been carried out to increase traffic safety on roads through facial recognition of drivers and the detection of signs of drowsiness. From this analysis, the optimal design to meet the established objective has been studied. This document presents a study of Machine Learning techniques and face detection tools applied to a program for facial recognition and detection of drowsiness in drivers.

Keywords: traffic safety, facial recognition, drowsiness detection, biometrics

Índice General

Capítulo 1: Introducción	9
Capítulo 2: Reconocimiento Facial de Conductores.....	11
Capítulo 3: Detección de Signos de Somnolencia	17
Capítulo 4: Inmovilización del Vehículo y Envío de Alerta con CAN Bus ...	20
Capítulo 5: Arquitectura	21
Capítulo 6: Herramientas	23
Capítulo 7: Conclusiones y Trabajo Futuro	25
Capítulo 8: Conclusions and Future Work.....	26
Bibliografía	27
Anexos	29

Índice de figuras

Ilustración 1: Flujo del funcionamiento del reconocimiento facial.....	12
Ilustración 2: HOG para detección de rostros	13
Ilustración 3: Landmark facial.....	14
Ilustración 4: Aplicación del filtro a la imagen.....	15
Ilustración 5: Arquitectura de una CNN.....	16
Ilustración 6: Detección de rostro marcando las regiones de interés	18
Ilustración 7: Flujo de detección de párpados abiertos o cerrados	18
Ilustración 8: Envío de alerta por detección de signos de somnolencia	19
Ilustración 9: Arquitectura del proyecto	21
Ilustración 10: Envío señal Ok / Not Ok al CAN Bus del vehículo	22
Ilustración 11: Herramientas analizadas	23
Ilustración 12: Diseño de la arquitectura de la base de datos	24

Capítulo 1: Introducción

Gracias a sucesivas mejoras en las tecnologías automovilísticas, y a las modificaciones en las normas de seguridad vial que han implicado el recrudecimiento de las sanciones por algunos delitos viales, se ha conseguido disminuir considerablemente el número de víctimas de accidentes de tráfico en los últimos años. Sin embargo, mientras sigan existiendo fallecimientos en accidentes viales, sigue siendo necesario seguir tomando más y mejores medidas de prevención.

Entre las causas más frecuentes de los accidentes de tráfico se encuentra el exceso de velocidad, seguido del consumo de alcohol y drogas, que causan más del 60\% de los accidentes al volante. Por otro lado, entre el 15\% y el 30\% de los accidentes está causado, directa o indirectamente, por la somnolencia de los conductores. Es ese un trastorno del sueño que conlleva una repentina e intensa necesidad de dormir, que conduce a quedarse dormido sin poder evitarlo [1].

Existe bastante bibliografía reciente relacionada con el reconocimiento del cansancio al volante [2] [3], pero dada la cantidad de consideraciones que se deben tener en cuenta en sistemas reales por la heterogeneidad de situaciones, sigue siendo de gran interés su estudio para aplicaciones concretas, como los sistemas de prevención anti-arranque.

El objetivo de este trabajo es diseñar una herramienta con dos funciones diferenciadas. Por un lado, el reconocimiento facial de los conductores permitirá acceder a consultas de información de tráfico relativa a la legalidad del permiso de conducir o a los delitos cometidos al volante por la persona identificada, impidiendo de esa forma el arranque del vehículo si así lo indican los datos recolectados. Por otro lado, la pronta detección de signos de somnolencia en el conductor permitirá alertar de forma temprana de que existe una alta probabilidad de que dicha persona se quede dormido al volante, evitando así que pueda suceder un accidente.

Para poder llevar a cabo el objetivo establecido, se hace necesario detectar los rostros de los conductores en las imágenes, de manera que se pueda averiguar tanto de quién se trata como si se está quedando dormido o no. Con esa meta, en el diseño propuesto en este trabajo se aplican varias técnicas de *Machine Learning* como el Histograma de Gradientes Orientados (HOG, *Histogram of Oriented Gradients*) para la detección de rostros [4] y la Estimación de *Landmark* Facial [5]. Estas técnicas permiten estimar la orientación del rostro para centrarlo al máximo posible, con objeto de poder luego identificar de quién se trata, comparando el *embedding* del conductor, generado con una Red Neuronal Convolutiva, con los *embeddings* de los conductores registrados [6].

Después de haber detectado el rostro, se requiere, en primer lugar, comprobar si la persona tiene los párpados abiertos o cerrados, haciendo uso también de una Red Neuronal Convolutiva, que permite clasificar en qué estado se encuentran. Después, se estudia si el conductor está bostezando o no y cuántas veces lo hace, haciendo uso del Modelo de Reconocimiento de Actividad Humana [7].

Capítulo 2: Reconocimiento Facial de Conductores

Una de las maneras que se proponen en este trabajo para intentar disminuir los accidentes viales es mediante el uso de algunos datos de tráfico del conductor, como posesión de licencia de conducir, tenencia de todos los puntos, infracciones cometidas, multas y si están pagadas o no, contratación de seguro, información del vehículo, etc. De esta manera, si el usuario al volante no cumple los requisitos necesarios para conducir el vehículo, se pone en marcha un mecanismo que inutiliza el motor, enviando una señal al CAN (*Controller Area Network*) Bus del vehículo [8], dando así lugar a que no se pueda conducir ese vehículo en ese momento.

El flujo de funcionamiento se muestra en la Ilustración 1. En primer lugar, se detectan las caras existentes en el vídeo que se captura a través de la cámara que se encuentra en el vehículo. Después el programa calcula el *embedding* de esa cara, es decir, los datos que permiten diferenciarla de cualquier otra cara y que le dan esa característica de unicidad. Luego se compara ese *embedding* con los *embeddings* previamente almacenados en la base de datos, es decir, se hace una comparación con las caras ya conocidas. Para la construcción de esa base de datos se propone que a la hora de obtener o renovar la licencia de conducir, se realicen varias fotografías del conductor que permitan calcular el *embedding* de su cara para almacenarlo de forma segura junto con su información de tráfico. Finalmente, una vez reconocido el conductor, se analizan sus datos de tráfico y si no es apto para conducir ese vehículo se envía señal al inmovilizador del motor.

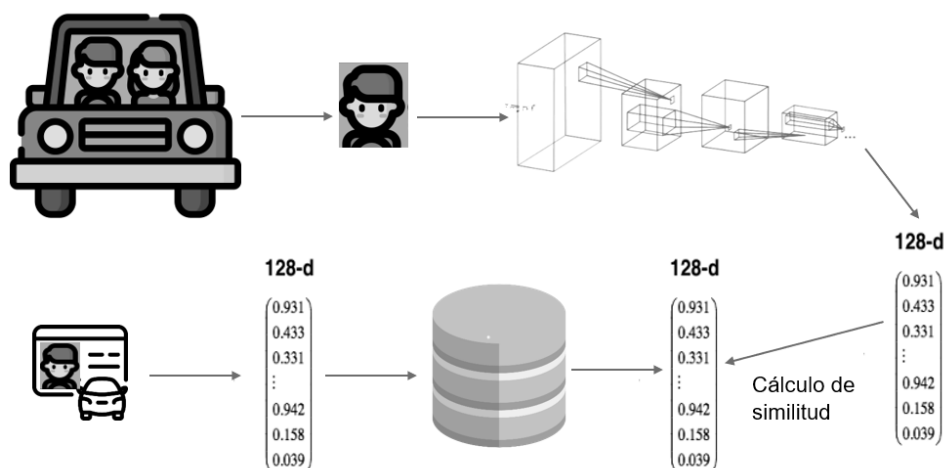


Ilustración 1: Flujo del funcionamiento del reconocimiento facial

Para llevar a cabo todo el proceso del reconocimiento facial son necesarios los siguientes pasos:

2.1: Detección de Rostros con Histograma de Gradientes Orientados

El primer paso para el reconocimiento facial es la detección de los rostros presentes en una imagen. Para esto, se utiliza el Histograma de Gradientes Orientados [4], que es un tipo de descriptor de características que se encarga de representar los rasgos elementales de un objeto como la forma, el color o la textura. Este mecanismo funciona observando cada píxel que conforma la imagen y sus píxeles vecinos. El objetivo del HOG es calcular cuán oscuro es el píxel actual comparado con los píxeles vecinos. Para ello, se dibujan flechas mostrando la dirección en la que la imagen se hace más oscura. Estas flechas, llamadas gradientes, se utilizan para mostrar el flujo de tonos claros a oscuros en la imagen. Finalmente, con la repetición de este proceso se consigue convertir cada uno de los píxeles de la imagen en una flecha.

De esa manera se consigue extraer la estructura básica del objeto contenido en la imagen. Sin embargo, al realizar el proceso descrito, se obtienen como resultado muchísimas flechas que serían difíciles de evaluar de manera rápida y eficiente. Es por esto por lo que la imagen se divide entonces en pequeños cuadrados de píxeles y en cada uno de ellos se cuenta cuántos gradientes

apuntan en cada dirección principal, de manera que después se reemplaza cada cuadrado en la imagen con las direcciones más fuertes y visibles. De esta forma, se consigue convertir la imagen original en una representación de la estructura básica del objeto, en este caso, un rostro (ver Ilustración 2). Entonces para descubrir si ese objeto se corresponde con un rostro, basta con comparar esa nueva imagen HOG con un patrón HOG ya conocido, calculado gracias a una gran cantidad de rostros de entrenamiento.



Ilustración 2: HOG para detección de rostros

2.2: Proyección de Rostros con Estimación de Landmark Facial

Una vez detectado un rostro en una imagen, se debe tener en cuenta que no siempre esos rostros están de frente a la cámara, sino que pueden estar de perfil o en otras posiciones. Por eso se debe posicionar el rostro capturado de manera que sea más fácil de estudiar y comparar por el modelo. Para ello, se hace uso de un algoritmo denominado Estimación de *Landmark* Facial [9], que tiene como idea principal describir un rostro a través de 68 puntos específicos (ver Ilustración 3), que es lo que se denomina como *landmark* y que se puede encontrar en cualquier rostro. De esta manera, una vez se sabe dónde están ubicados cada uno de los puntos que detallan dónde se hallan los ojos, la nariz,

la boca y demás partes que conforman un rostro, se hace una rotación y escalado, de forma que los ojos y la boca estén lo más centrados posibles, para hacer más sencillo el proceso de cálculo del *embedding* facial.

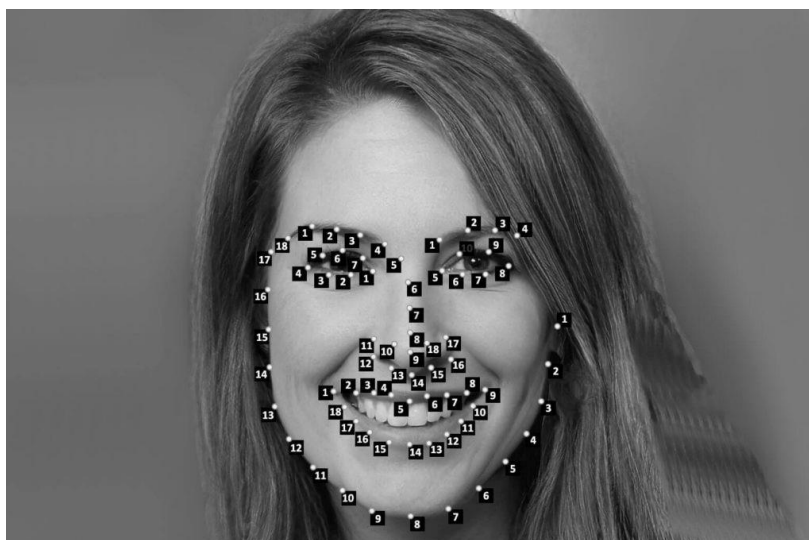


Ilustración 3: Landmark facial (Fuente: www.latesttechnicalreviews.com)

2.3: Generación de Embedding Facial con Redes Neuronales Convolucionales

El siguiente paso consiste en calcular el *embedding* del rostro detectado en la imagen, con el fin de compararlo con los demás *embeddings* almacenados en la base de datos de los rostros ya conocidos. Estos *embeddings* son, realmente, 128 números que representan las características únicas y más importantes de un rostro [6] y para calcularlos es necesario hacer uso de una Red Neuronal Convolutiva (CNN, *Convolutional Neural Network*), que es un tipo de red neuronal artificial cuyo diseño ha sido pensado para sacar partido a la estructura espacial de una imagen, ya que tiene la capacidad de interpretar las formas y patrones más complejos presentes en grandes conjuntos de imágenes. Las neuronas de las CNNs son muy similares a las neuronas en la corteza visual primaria de un cerebro biológico, es decir, su procedimiento de aprendizaje se asemeja mucho al proceso de visión de un ser humano.

Siguen un procesamiento mediante el cual se identifican primero elementos básicos y generales, que luego son combinados con el fin de generar patrones cada vez más complejos.

Este tipo de redes neuronales se caracterizan por realizar una serie de convoluciones. En procesamiento de imágenes, una convolución no es más que una operación matemática que, realizando combinaciones con los valores de los píxeles, es capaz de generar nuevas imágenes con las que estudiar formas y patrones que componen a la imagen original. Cada píxel nuevo que se vaya a generar se calcula aplicando una matriz de números, llamada filtro o *kernel*, sobre la imagen original, y luego se multiplican y suman los valores de cada píxel vecino para obtener así el nuevo valor (ver Ilustración 4). Por tanto, la principal característica de una CNN es aplicar los filtros necesarios para detectar todas las formas y patrones que conforman la imagen. Cada vez que se utiliza un filtro, se genera una nueva imagen, denominada mapa de características, que actúa como un mapa donde se indica en qué parte de la imagen se ha detectado la característica buscada por el filtro.

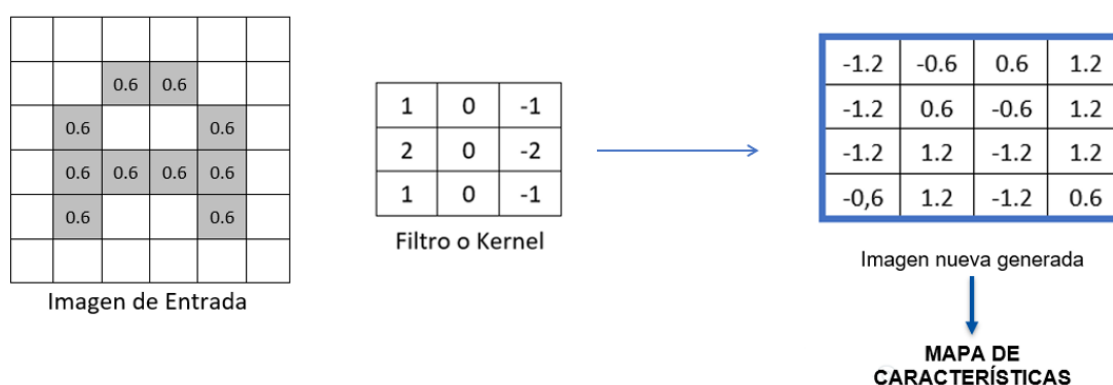


Ilustración 4: Aplicación del filtro a la imagen

Lo destacable de la CNN es que el proceso de búsqueda de características se realiza de forma secuencial, es decir, la salida de una de las capas se convierte en la entrada de la siguiente. Por eso la operación de convolución se hace cada vez más potente, puesto que cada vez que se realiza una convolución se está

accediendo a más información espacial de la imagen original. Con las convoluciones lo que se está haciendo son nuevas detecciones sobre las detecciones recibidas de capas anteriores.

En otras palabras: realizar las detecciones sobre detecciones anteriores permite componer cada vez patrones más complejos.

La arquitectura de una CNN (ver Ilustración 5) se representa normalmente como un embudo, ya que la imagen inicial se va comprimiendo espacialmente, es decir, su resolución va disminuyendo, al mismo tiempo que su grosor va en aumento, puesto que el número de mapas de características que se detectan crece. Finalmente, se consigue obtener todos los patrones necesarios de la imagen, que luego se introducen como *inputs* independientes dentro de una red neuronal multicapa para descubrir qué representa esa imagen. En el caso que nos ocupa, calcula el *embedding* del rostro que se esté analizando.

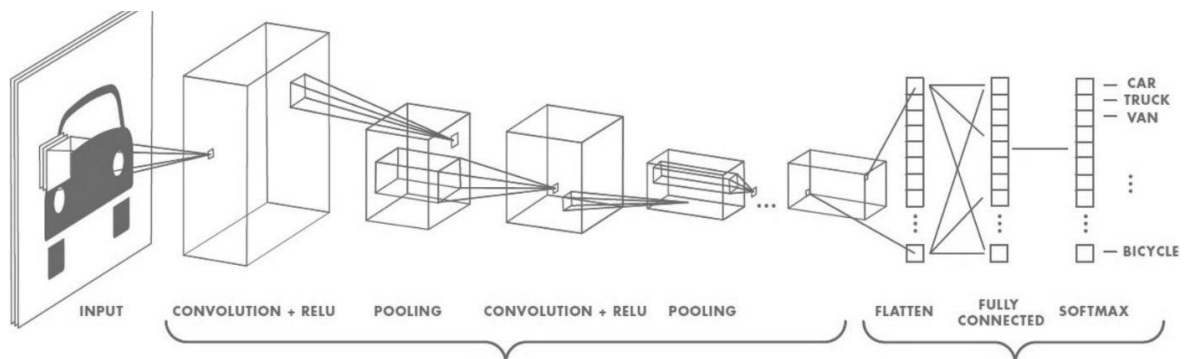


Ilustración 5: Arquitectura de una CNN (Fuente: www.es.mathworks.com/discovery/convolution.html)

Capítulo 3: Detección de Signos de Somnolencia

Otra propuesta para intentar disminuir el número de accidentes de tráfico y víctimas mortales es detectando de manera temprana los signos de somnolencia que pudiera estar presentando el conductor, con el fin de poner en marcha una alerta que le haga despertarse y tomar conciencia de su estado.

En la línea de varios trabajos previos, como [10] o [11], usaremos redes neuronales convoluciones para la detección de la somnolencia.

Para realizar este proceso, se estudian las imágenes recibidas a través de la cámara del vehículo, detectando rostros, ojos y boca, para observar el comportamiento de estas partes del cuerpo y determinar si están pasando por un proceso de somnolencia, ya sea por la frecuencia de parpadeo, por la regularidad con la que suceden los bostezos en un periodo de tiempo determinado o por movimientos involuntarios de cabeza que indiquen que una persona se está dejando dormir.

Este cometido se puede cumplir llevando a cabo los siguientes pasos: primero, se toman las imágenes capturadas por la cámara de vídeo y se detecta el rostro del conductor como se muestra en la Ilustración 6, el cual se marca como una Región de Interés (ROI, *Region Of Interest*). Dentro de esa ROI se detectan los ojos con los que se alimenta el clasificador, que en este caso es también una CNN. Ésta se encarga de comprobar si los párpados del conductor están abiertos o cerrados (ver Ilustración 7).

Más adelante, se detecta si la persona que conduce el vehículo está bostezando o no. Para ello, se hace uso del Modelo de Reconocimiento de Actividad Humana [12] con *OpenCV*, que es una biblioteca libre de visión artificial [13].

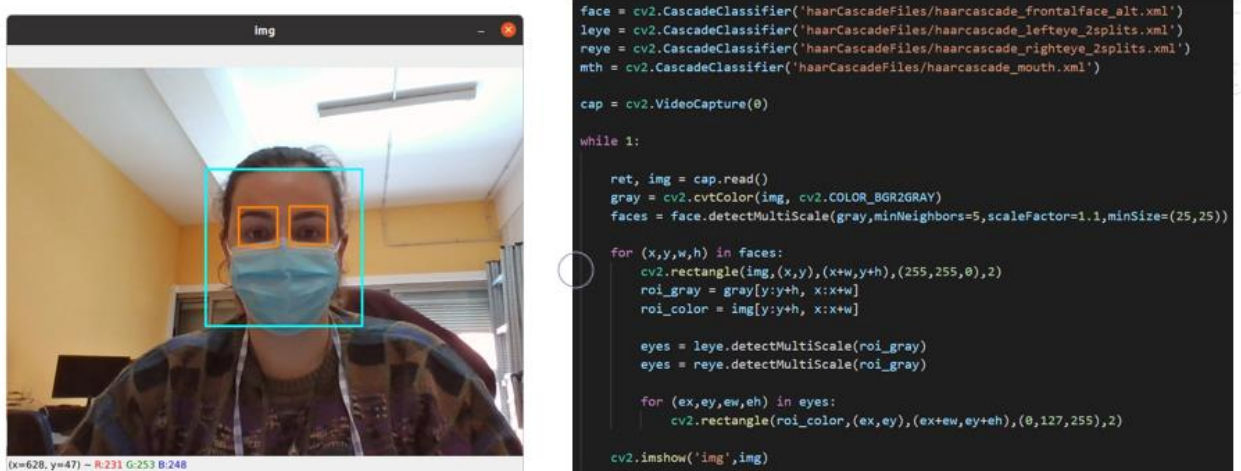


Ilustración 6: Detección de rostro marcando las regiones de interés

Dicho modelo ha sido previamente entrenado con el *Dataset* de Vídeos de Acciones Humanas de *Kinetics* [14] de forma que detecta si el conductor está bostezando y cuántas veces lo hace. En este modelo, se han utilizado arquitecturas 2D existentes que han sido extendidas a través de *kernel*s 3D para llevar a cabo la clasificación de vídeos.

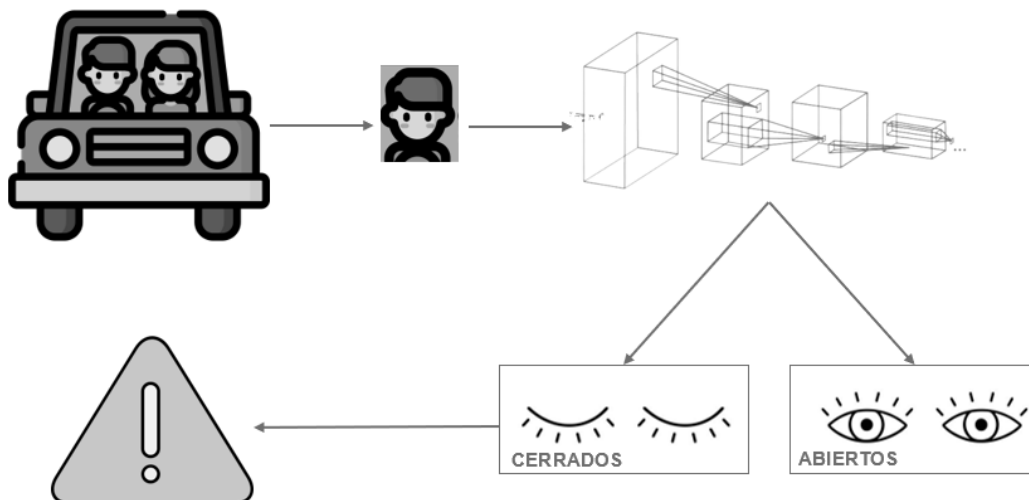


Ilustración 7: Flujo de detección de párpados abiertos o cerrados

Una vez clasificados estos datos, si se comprueba que el conductor tiene los párpados cerrados durante más segundos de lo normalmente establecido, y que presenta una frecuencia de bostezo bastante alta, se lanza una señal a una alarma conectada al CAN Bus del vehículo, con el fin de alertar al conductor de que se está quedando dormido o de que tiene una alta probabilidad de hacerlo (ver Ilustración 8).

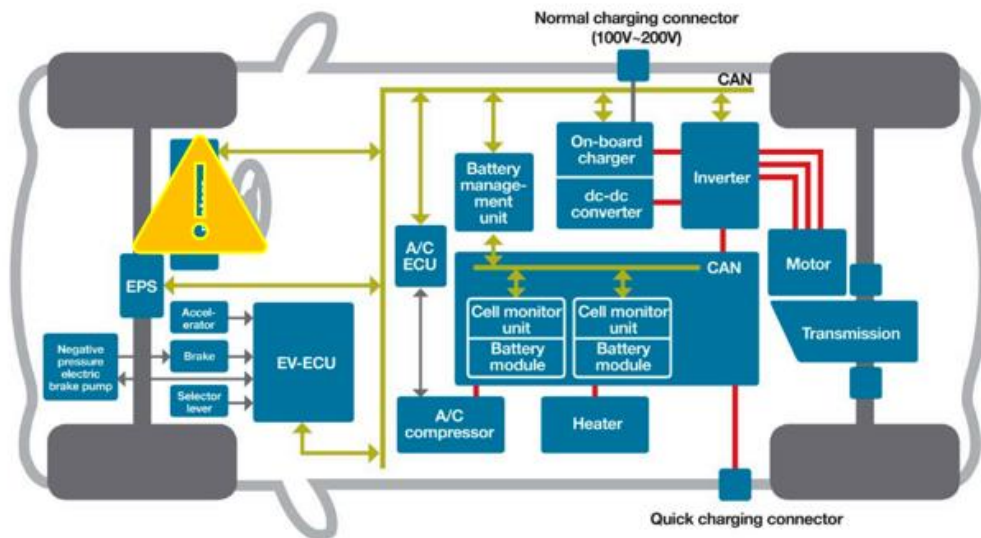


Ilustración 8: Envío de alerta por detección de signos de somnolencia

Capítulo 4: Inmovilización del Vehículo y Envío de Alerta con CAN Bus

El CAN Bus es un sistema diseñado por la compañía alemana *Robert Bosch GmbH*, que fue lanzado oficialmente en 1986 en el congreso de la Sociedad de Ingenieros Automotrices (SAE, *Society of Automotive Engineers*). Este sistema en serie basado en una topología bus permite intercambiar información entre los distintos componentes electrónicos de un vehículo, como son, por ejemplo, el motor, los asientos eléctricos, el climatizador, el techo solar, etc.

Por una parte, durante el arranque del vehículo, una vez se ha identificado al conductor mediante reconocimiento facial, se accede a sus datos de tráfico con el fin de saber si está en condiciones de conducir ese vehículo o no. Para esto, la cámara a través de la cual se capturan las imágenes debe estar conectada al CAN Bus del vehículo, de forma que cuando se consiga reconocer al conductor y se sepa si puede conducir el vehículo, se envíe una señal al inmovilizador del motor indicando si puede ser arrancado o no.

Por otra parte, durante la conducción, en el supuesto de que se detecten signos de somnolencia en el conductor, se enviará una señal de alarma también a través del CAN Bus del vehículo, ya sea mediante un mensaje visual en la pantalla integrada, un mensaje sonoro o bien la mezcla de ambos.

Capítulo 5: Arquitectura

La arquitectura de este proyecto (ver Ilustración 9) ha sido diseñada para proveer al vehículo de la información necesaria del conductor, es decir, para registrar la llave digital del conductor en el vehículo. Por tanto, cuando una persona realice la compra de un vehículo, el gerente del concesionario cargará la llave digital del comprador en el vehículo en cuestión. Es decir, se hará una conexión a la base de datos de tráfico con la que se traigan los datos necesarios del conductor, que en este caso concreto será el *embedding* facial y, una vez recolectado, se almacenarán en el dispositivo que se encuentre instalado en el vehículo.

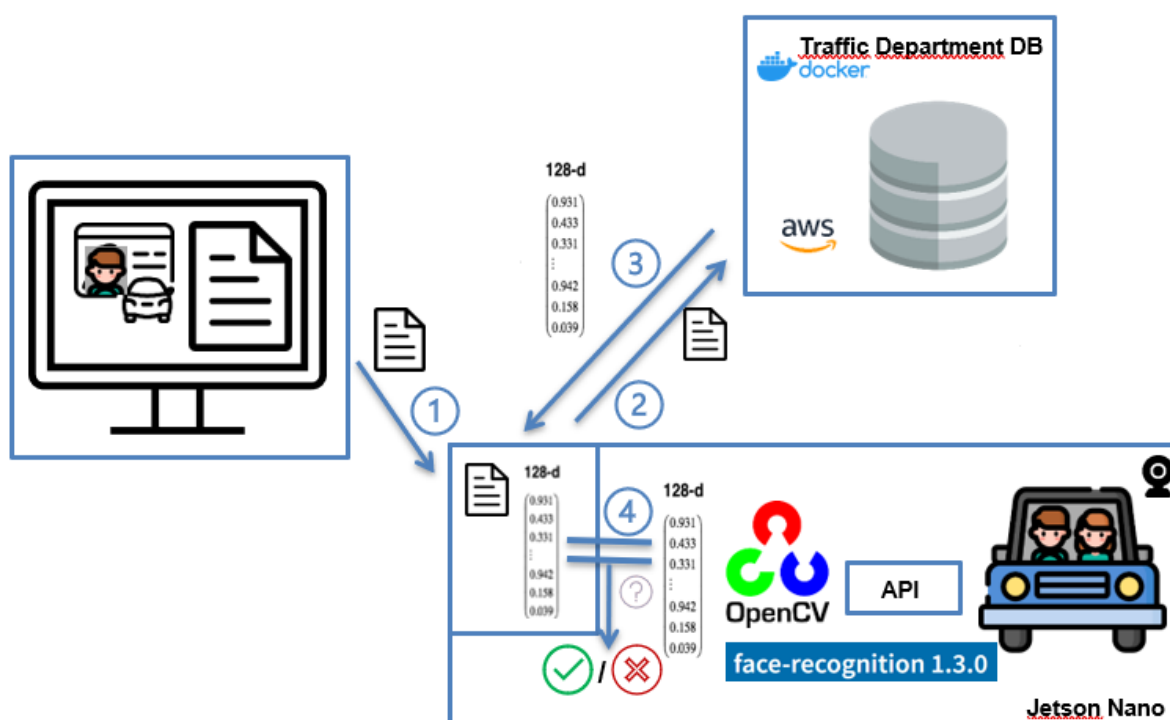


Ilustración 9: Arquitectura del proyecto

De esta manera, cuando una persona desee poner en marcha el vehículo se seguirán los siguientes pasos:

1. Comparación del *embedding* facial almacenado en el dispositivo del vehículo con el calculado en *streaming* a través de la cámara situada en el interior.
2. Si se comprueba que se trata de alguna de las personas con permiso para conducir el vehículo, se accede a la base de datos para comprobar su información de conducir: permiso en regla, puntos, infracciones.
3. Si se confirma que está todo correcto, se envía una señal de *Ok* al CAN Bus del vehículo, que permitirá el arranque del motor. En el caso contrario, se enviará una señal de *Not Ok*, con la que no será posible la puesta en marcha del vehículo.

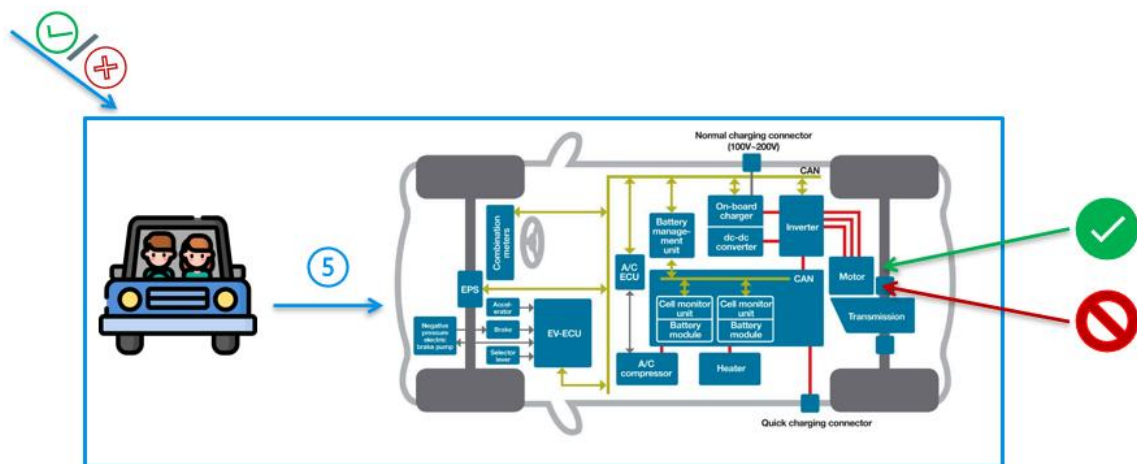


Ilustración 10: Envío señal Ok / Not Ok al CAN Bus del vehículo

4. La actualización del *embedding* facial se hará cuando el vehículo sea provisto de la llave digital y cuando se renueve el carné de conducir, de forma que se tenga la versión facial más actualizada de cada conductor.

Capítulo 6: Herramientas

Para diseñar el sistema con todos los pasos comentados, previamente se estudiaron las herramientas existentes en el mercado con funcionalidades similares.

Entre todas las herramientas analizadas (ver Ilustración 11) destaca la librería de *Python*, *face_recognition* [15], que permite reconocer y manipular rostros. Dicha librería está creada con *Dlib*, un *kit* de herramientas que contiene algoritmos de *Machine Learning* y mecanismos para crear softwares complejos. Fue creada por Adam Geitgey en 2017, consiguiendo encapsular en una sola librería todos los pasos descritos en el [Capítulo 2](#), con el fin de trabajar con rostros capturados en imágenes.



Ilustración 11: Herramientas analizadas

Para ambas fases de este trabajo se puede utilizar dicha librería, pues facilita mucho el trabajo y permite realizar el reconocimiento y la detección de rostros de una manera rápida y eficiente. La captura de las imágenes de la cámara de vídeo se hace gracias al uso de *OpenCV* y *Machine Learning*. Con *OpenCV* no solo se puede realizar la detección de rostros, sino que también se pueden

detectar partes que conforman el rostro como los ojos, la nariz o la boca, además de clasificar acciones humanas captadas en vídeo, como, por ejemplo, la acción de bostezar.

A la hora de acceder a los datos de tráfico de un conductor, haría falta simular la base de datos de la Dirección General de Tráfico. Por ello, en principio, se podría hacer uso de una base de datos no relacional como *MongoDB*, ya que lo natural es que crezca cada vez más, aumentando con rapidez su tamaño considerablemente. Un posible diseño de la arquitectura de la base de datos es el que se muestra en la Ilustración 12, en la que se puede ver una tabla para recoger la información de los conductores, otra para los datos de sus licencias de conducir y una final para almacenar los vehículos con toda su información.

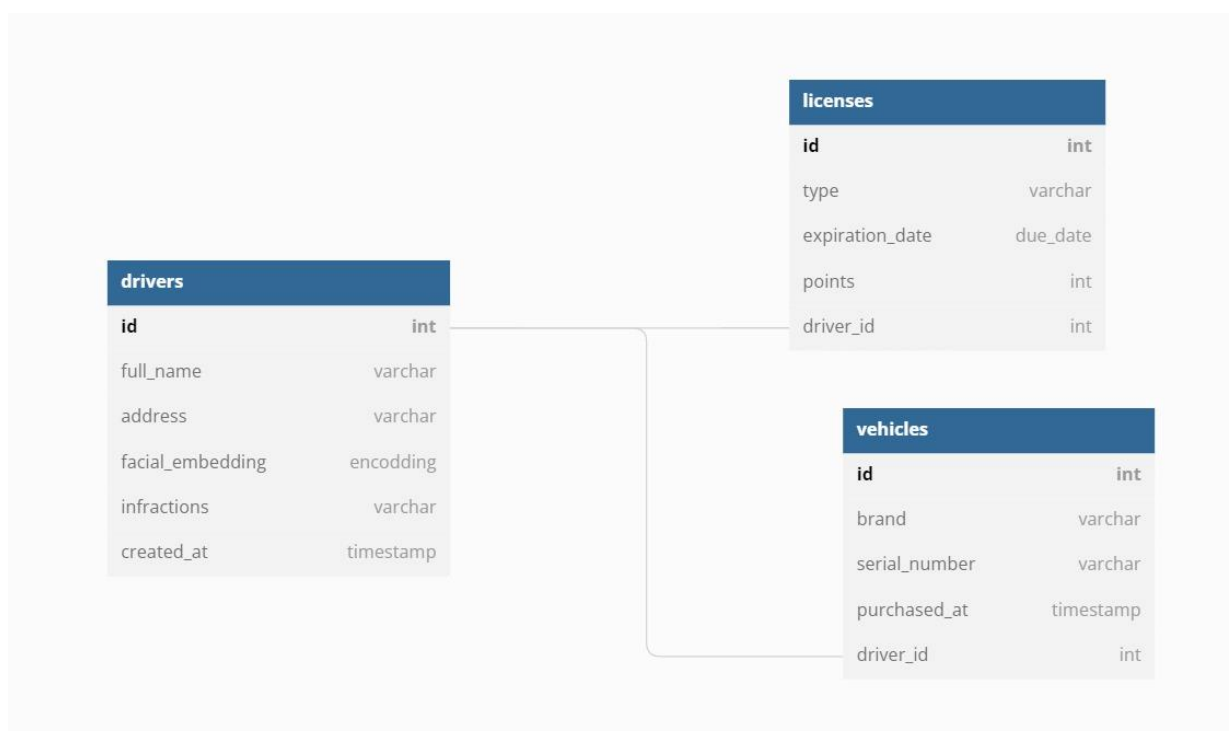


Ilustración 12: Diseño de la arquitectura de la base de datos

La API con la que se accede a la base de datos se haría con *FastAPI*, el cual es un *framework* de *Python* para generar APIs web. En relación con la protección de la información almacenada, se tendrán en cuenta las últimas consideraciones de seguridad relativas a bases de datos NoSQL [16].

Capítulo 7: Conclusiones y Trabajo Futuro

El número de víctimas mortales en las carreteras sigue siendo alarmante. Gran parte de estos accidentes son ocasionados por somnolencia. El principal objetivo de este trabajo es proporcionar una contribución en este sentido para hacer más seguras las carreteras, de manera que se consiga reducir el porcentaje de accidentes viales producidos con ese motivo.

Para ello, se ha realizado un estudio de algunas de las herramientas existentes en el mercado, de cómo funcionan y de cómo se pueden utilizar para cumplir el objetivo propuesto. Para conseguir el cometido, se ha realizado la búsqueda de algunas herramientas existentes, intentando elegir las más adecuadas y eficientes, con las que, mediante su combinación, se pueda dar lugar a un programa capaz de cumplir con los requisitos establecidos en este proyecto: por un lado, el reconocimiento facial del conductor que permite comprobar sus datos de tráfico y ver si posee las condiciones adecuadas para conducir. Por otro lado, la detección temprana de signos de somnolencia en el conductor para alertar de manera temprana la posibilidad de quedarse dormido al volante.

Por tanto, después de haber realizado un estudio del mercado y de haber elegido los instrumentos necesarios, se ha llevado a cabo el diseño del modelo contemplando la implementación tanto del reconocimiento facial de conductores como de la detección de signos de somnolencia, de manera que en un futuro se pueda llegar a hacer una demostración en vivo de la utilidad de la herramienta diseñada. Ya que se trata de un trabajo actualmente en desarrollo, se hace necesaria una implementación optimizada, así como la realización de tests simulados y en vivo que permitan comprobar el correcto comportamiento de las funcionalidades desarrolladas, así como su actualización y mejora.

Capítulo 8: Conclusions and Future Work

The death toll on the roads remains alarming. Many of these accidents are caused by drowsiness. The main objective of this work is to provide a contribution in this sense to make the roads safer, so that the percentage of road accidents produced by this cause is reduced.

To this end, a study has been made of some of the existing tools on the market, how they work and how they can be used to fulfill the proposed objective. In this sense, although there are several tools that can be used to achieve the mission, in this work it has been decided to investigate such tools, choose the most appropriate and efficient, and design one of its own from the information obtained and the means available.

In summary, what has been done in this work is to combine some existing tools, conveniently chosen, to produce a program capable of meeting the requirements established to achieve the proposed objective. On the one hand, driver facial recognition allows to check driver traffic data and see if they have the right conditions to drive. On the other hand, early detection of signs of drowsiness in the driver allows early warning of falling asleep behind the wheel.

Therefore, after having carried out a study of the market and having chosen the necessary tools, the design of the model has been carried out considering the implementation of both the facial recognition of drivers and the detection of signs of drowsiness, so that a live demonstration of the utility of the designed tool can be made in the future. Bearing in mind that this is a work currently under development, an optimized implementation is necessary as well as the realization of simulated and live tests that allow checking the correct behavior of the functionalities developed, as well as its updating and improvement.

Bibliografía

- [1] «Conducir con sueño o cansancio,» 2022. [En línea]. Available: <https://www.dgt.es/muevete-con-seguridad/evita-conductas-de-riesgo/Conducir-con-sueno-o-cansancio>.
- [2] Y. G. H. T. J. L. a. H. W. Feng You, «A fatigue driving detection algorithm based on facial motion information entropy,» *Journal of Advanced Transportation*, 2020.
- [3] Y. M. T. a. M. A. Albadawi, «A review of recent developments in driver drowsiness detection systems,» *Sensors*, vol. 22(5), p. 2069, 2022.
- [4] N. D. a. B. Triggs, «Histograms of Oriented Gradients for Human Detection,» de *EEE Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005.
- [5] Y. W. a. Q. Ji, «Facial landmark detection: A literature survey,» *International Journal of Computer Vision*, vol. 127(2), pp. 115-142, 2019.
- [6] D. K. a. J. P. Florian Schroff, «FaceNet: A Unified Embedding for Face Recognition and Clustering,» de *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [7] Z. W. a. H. W. Lin Fan, «Human activity recognition model based on decision tree,» de *IEEE International Conference on Advanced Cloud and Big Data*, 2013.
- [8] I. 11898-1:2015, «Road vehicles — Controller area network (CAN) — Part 1: Data link layer and physical signalling,» [En línea]. Available: <https://www.iso.org/standard/63648.html>.
- [9] V. K. a. J. Sullivan, «One Millisecond Face Alignment with an Ensemble of Regression Trees,» de *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [10] M. N.-M. G. S.-P. a. H. P.-M. Jonathan Flores-Monroy, «Visual-based Real Time Driver Drowsiness Detection System Using CNN,» de *18th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, 2021.
- [11] S. J. a. M. F. Abid Ali Minhas, «A smart analysis of driver fatigue and drowsiness detection using convolutional neural networks,» *Multimedia Tools and Applications*, pp. 05-06, 2022.

- [12] K. H. e. al., «Can Spatiotemporal 3D CNNs Retrace the History of 2D CNNs and ImageNet?,» *arXiv from Cornell University*, 2017.
- [13] «OpenCV,» [En línea]. Available: <https://opencv.org>.
- [14] W. K. e. al., «"The Kinetics Human Action Video Dataset,» *arXiv from Cornell University*, 2017.
- [15] A. Geitgey, «Face Recognition Documentation,» 2019. [En línea]. Available: <https://face-recognition.readthedocs.io/en/latest/readme.html>.
- [16] A. R. a. A. C.-P. Sabrina Sicari, «Security & privacy issues and challenges in NoSQL databases,» de *Computer Netwroks*, ELSEVIER, 2022, p. Vol. 206.

Anexos

Anexo I: Artículos enviados a congresos

Drivers Facial Recognition and Drowsiness Detection

Alba Cruz-Torres, Pino Caballero-Gil, Candelaria Hernández-Goya.

Workshop on Cybersecurity Applications and Intelligent Transportations Systems

CAITIS.

Proceedings of the International Conference on Security and Management SAM

World Congress in Computer Science, Computer Engineering, and Applied Computing

CSCE

Las Vegas, USA. July 25-28, 2022.

Springer Nature.

Indexada en Computing Research and Education (CORE), con ranking C

Indexada en CS Conference Rankings (0.83)

Indexada en GII-GRIN en Class WiP

Reconocimiento Facial e Identificación de Somnolencia en Conductores

Alba Cruz-Torres, Carlos Rosa-Remedios, Pino Caballero-Gil, Candelaria Hernández-Goya.

XVII Reunión Española sobre Criptología y Seguridad de la Información RECSI.

Santander, 19-21 octubre 2022