

Grado Universitario en Ingeniería Informática
2020-2021

Trabajo Fin de Grado

“Reconocimiento de Emociones en los Acompañantes de un Vehículo”

Binxian Huang

Tutor/es

Jose Antonio Iglesias Martínez
Leganés, Junio 2024



Esta obra se encuentra sujeta a la licencia Creative Commons **Reconocimiento - No Comercial - Sin Obra Derivada**

RESUMEN

El presente trabajo de fin de grado se centra en el desarrollo de un modelo de reconocimiento de emociones faciales para su aplicación en sistemas de vehículos, con el objetivo de mejorar la experiencia del usuario y aumentar la seguridad en la conducción. Aprovechando los avances en inteligencia artificial y aprendizaje automático, se propone un sistema capaz de identificar en tiempo real las emociones de los pasajeros y ajustar de manera dinámica las configuraciones internas del vehículo, tales como música, iluminación y temperatura, en función del estado emocional detectado.

El proyecto aborda varias etapas clave: análisis del estado del arte, diseño del sistema, implementación del modelo, y evaluación de resultados. En el análisis del estado del arte se revisan las tecnologías actuales en la industria automotriz, incluyendo los Sistemas Avanzados de Ayuda a la Conducción (ADAS) y los vehículos autónomos, destacando la importancia de integrar la IA para una experiencia de usuario personalizada.

El diseño del sistema incluye la identificación de requisitos funcionales y no funcionales, y la selección de herramientas y técnicas adecuadas para el procesamiento y análisis de datos. Para la implementación, se utilizó el dataset FER+ y se emplearon técnicas de preprocesamiento y redes neuronales convolucionales (CNN) para entrenar el modelo de reconocimiento emocional. Los resultados obtenidos indican una precisión razonable del modelo, con un porcentaje de aciertos del 70.14 %, concluyendo como posible objetivo futuro una mejora en la calidad de las imágenes y el preprocesamiento para aumentar la precisión.

Finalmente, se discuten los impactos socio-económicos y regulatorios del proyecto, subrayando la necesidad de cumplir con las normativas de protección de datos y garantizar la privacidad y seguridad de los datos personales. Este proyecto representa un avance significativo en la personalización y seguridad de los sistemas automotrices, aportando beneficios tanto para los usuarios como para la industria.

Palabras clave: Reconocimiento de emociones, Vehículo, Pasajeros, Inteligencia Artificial, Aprendizaje Automático, Vehículos Inteligentes, Experiencia del Usuario, Seguridad en la Conducción, Sistemas Avanzados de Ayuda a la Conducción (ADAS), Redes Neuronales Convolucionales (CNN).

DEDICATORIA

Quiero agradecer a mis padres, por su amor y apoyo en cada momento de mi vida, y por estar allí siempre en todo momento.

A mis amigos y compañeros, por brindarme su compañía y alegría en los momentos difíciles y en los más felices.

A mis profesores, en especial a mi tutor Jose Antonio Iglesias Martinez, por su paciencia y orientación en la realización de este trabajo.

Y finalmente, a todas aquellas personas que, de una forma u otra, han contribuido a mi crecimiento personal y académico.

ÍNDICE GENERAL

1. INTRODUCCIÓN	1
1.1. Motivación	1
1.2. Objetivos	1
1.3. Estructura de la memoria	2
2. ESTADO DEL ARTE	3
2.1. Experiencia del Usuario en la Industria de la Automoción	3
2.1.1. Paneles táctiles	3
2.1.2. Coches conectados	5
2.1.3. Heads Up Display (HUD)	8
2.2. Sistemas Avanzados de Ayuda a la Conducción (ADAS)	10
2.3. Vehículos Autónomos	16
2.4. Inteligencia Artificial y Aprendizaje Automático	20
2.4.1. Redes Neuronales	21
2.5. Trabajos Relacionados	24
2.6. Conclusiones del Estado del Arte	25
3. DISEÑO DEL SISTEMA	27
3.1. Análisis de requisitos	28
3.1.1. Requisitos funcionales	29
3.1.2. Requisitos no funcionales	32
3.2. Análisis de datos	34
3.3. Análisis de sistemas y software utilizado	36
4. IMPLEMENTACIÓN Y RESULTADOS	37
4.1. Preprocesado de datos	37
4.2. Entrenamiento del modelo	39
4.3. Análisis de resultados	42
5. MARCO REGULADOR E IMPACTO SOCIO-ECONÓMICO	45
5.1. Marco regulador	45
5.2. Impacto socio-económico	46

6. GESTIÓN DEL PROYECTO	47
6.1. Planificación	47
6.2. Estimación	48
6.2.1. Costes personal	48
6.2.2. Costes hardware	49
6.2.3. Costes software	50
6.2.4. Resumen de costes	50
7. CONCLUSIONES Y TRABAJOS FUTUROS	52
BIBLIOGRAFÍA	54

ÍNDICE DE FIGURAS

2.1	Evolución del panel en un Volkswagen Golf. Fuente: [1]	4
2.2	Conductor interaccionando con una pantalla táctil. Fuente: [1]	4
2.3	Ejemplos de ECUs en un coche. Fuente: [3]	5
2.4	Visión general de un sistema de coche conectado. Fuente: [3]	5
2.5	Ejemplos de interfaces de Android Auto en la pantalla del coche.	7
2.6	Ventas de coches con Apple CarPlay vs Android Auto. Fuente: [3]	8
2.7	Vista de un HUD de un Jaguar. Fuente: [4]	8
2.8	Funcionamiento y componentes de un HUD. Fuente: [4]	9
2.9	Visión de cámara de infrarrojos en el HUD. Fuente: [4]	9
2.10	Visión de HUD en espejos retrovisores. Fuente: [4]	10
2.11	Funcionamiento del sistema DDR. Fuente: [6]	11
2.12	Funcionamiento del sistema ISA. Fuente: [6]	12
2.13	Funcionamiento del sistema RCTA. Fuente: [6]	12
2.14	Funcionamiento del sistema EDR. Fuente: [6]	13
2.15	Funcionamiento del sistema LDW. Fuente: [6]	13
2.16	Funcionamiento del sistema BAS y EBA. Fuente: [6]	14
2.17	Funcionamiento del sistema Alcolock. Fuente: [6]	15
2.18	Funcionamiento del sistema de Alerta de cinturón en las plazas traseras. Fuente: [6]	16
2.19	Niveles de autonomía según el estándar SAE J3016. Fuente: [9]	17
2.20	Sensores en un vehículo autónomo. Fuente: [9]	18
2.21	Clasificación de sensores usados en sistemas ADAS. Fuente: [10] . . .	19
2.22	Rangos sobre los que actúan los sensores. Fuente: [10]	19
2.23	Campos de la Inteligencia Artificial. Fuente: [12]	20
2.24	Modelos de Aprendizaje Automático. Fuente: [12]	21
2.25	Artificial Neural Network basic architecture. Fuente: [13]	22
2.26	Recurrent Neural Network basic architecture. Fuente: [13]	22
2.27	Convolutional Neural Network basic architecture. Fuente: [13]	23

2.28 Convolutional Neural Network layers visualization. Fuente: [14]	24
3.1 Esquema de relación y funcionalidad de los módulos.	28
3.2 Ejemplos de reanotación de imágenes (arriba FER2013 original, abajo FER+). Fuente: [21]	35
3.3 Ejemplos de imágenes del dataset CK+.	35
3.4 Ejemplos de imágenes del dataset JAFFE.	36
4.1 Gráfico de distribución inicial de valores.	38
4.2 Gráfico de distribución balanceada de valores.	39
4.3 Gráficas de valores de función de pérdida y valores de precisión.	42
4.4 Matriz de confusión de predicciones del modelo.	43
6.1 Diagrama de Gantt. (Parte 1)	47
6.2 Diagrama de Gantt. (Parte 2)	48
6.3 Diagrama de Gantt. (Parte 3)	48

ÍNDICE DE TABLAS

3.1	Plantilla de tabla de requisitos	28
3.2	Requisito funcional RF-01	29
3.3	Requisito funcional RF-02	29
3.4	Requisito funcional RF-03	29
3.5	Requisito funcional RF-04	29
3.6	Requisito funcional RF-05	30
3.7	Requisito funcional RF-06	30
3.8	Requisito funcional RF-07	30
3.9	Requisito funcional RF-08	30
3.10	Requisito funcional RF-09	31
3.11	Requisito funcional RF-10	31
3.12	Requisito funcional RF-11	31
3.13	Requisito funcional RF-12	31
3.14	Requisito funcional RF-13	32
3.15	Requisito funcional RF-14	32
3.16	Requisito no funcional RNF-01	32
3.17	Requisito no funcional RNF-02	32
3.18	Requisito no funcional RNF-03	33
3.19	Requisito no funcional RNF-04	33
3.20	Requisito no funcional RNF-05	33
3.21	Requisito no funcional RNF-06	33
3.22	Requisito no funcional RNF-07	34
3.23	Tabla de comparación de datasets	35
4.1	Tabla de distribución inicial de valores	37
4.2	Tabla de distribución balanceada de valores	38
4.3	Tabla de comparación de diferentes datasets	39
6.1	Tabla de costes de personal	49

6.2	Tabla de costes hardware	50
6.3	Tabla de resumen de costes	51

1. INTRODUCCIÓN

1.1. Motivación

Actualmente, siendo una era de grandes avances tecnológicos, la industria automotriz destaca como un faro de innovación, particularmente en el desarrollo de sistemas inteligentes diseñados para mejorar la seguridad y personalizar la experiencia del conductor. La integración de la Inteligencia Artificial (IA) en los sistemas de los vehículos representan un cambio transformador, pasando de los vehículos tradicionales a los “vehículos inteligentes” o *smart vehicles*. Estos vehículos no son meros medios de transporte, sino que son plataformas sofisticadas que se adaptan de forma dinámica a las necesidades y estados emocionales de los pasajeros.

Las tecnologías en los vehículos actuales en gran medida no tienen en cuenta el contexto emocional de los pasajeros, dependen de una entrada en concreto como comandos táctiles o vocales. La implementación de sistemas capaces de percibir y reaccionar a los estados emocionales de los pasajeros podría ser un gran avance para el sector del diseño de software centrado en el usuario.

Este proyecto se centra en aprovechar el gran avance de la IA para desarrollar un modelo capaz de reconocer y analizar emociones faciales de los pasajeros en tiempo real, y aplicarlo en un sistema afectando a la forma en la que los vehículos interactúan con los ocupantes, ajustando configuraciones internas como la música, la iluminación y la temperatura, para mejorar el confort y la seguridad. Esto serviría mejorar la experiencia del usuario, además de que podría ser útil en los “Sistemas Avanzados de Ayuda a la Conducción” o Advanced Driver-Assistance System (ADAS), ya que ayudaría al conductor a librarse de realizar por sí mismo dichas acciones y no distraerse de la conducción, además que el modelo realizado también podría ser utilizado en los sistemas.

1.2. Objetivos

El objetivo principal de este proyecto es desarrollar un modelo de reconocimiento de emociones faciales, que pueda ser aplicado en el sistema de un vehículo para mejorar la experiencia del usuario. Para ello, se plantean los siguientes objetivos específicos:

- Desarrollar un modelo capaz de identificar las expresiones faciales mediante imágenes de las caras de los pasajeros.
- Diseñar un sistema que permita integrar el modelo con los demás sistemas del vehículo, además de poder interactuar con los controles del vehículo como la iluminación, los sistemas de audio o el control del clima.

- Mejorar las medidas de seguridad en función de los resultados del modelo, como permitir al sistema ajustar la dinámica de conducción, activar protocolos de seguridad, alertar sobre posibles peligros, etc. si se detectan señales de enfado o angustia.
- Análisis y estudio de viabilidad económica y técnica del sistema, verificando si la integración de esta tecnología es factible para el mercado comercial, además de un análisis de impacto y una evaluación de los usuarios en la aceptación de esta tecnología en los vehículos.

1.3. Estructura de la memoria

El documento se organiza en siete capítulos principales:

1. **Introducción:** En este capítulo se describe el contexto general del proyecto, incluyendo la motivación que impulsa a su realización y los objetivos específicos que se pretende alcanzar.
2. **Estado del Arte:** En este capítulo se presenta una breve revisión literaria de los temas relacionados al proyecto, los avances de la experiencia del usuario el sector de los automóviles, la inteligencia artificial en la industria de la automoción, los ADAS y los sistemas implementados que mejoran las experiencias de los usuarios, además de un análisis de los distintos trabajos previos en relación al tema.
3. **Diseño del sistema:** En este capítulo se describe el diseño y arquitectura del sistema, los requisitos necesarios, un análisis y explicación de la fuente de datos elegida para la realización del modelo, al igual que una descripción de las metodologías y herramientas que se usarán.
4. **Implementación y resultados:** En este capítulo se documenta el proceso práctico de la realización del modelo así como un análisis de los resultados obtenidos.
5. **Marco regulador e impacto socioeconómico:** En este capítulo se realiza en una primera sección una revisión de las normativas y regulaciones relevantes que afecten a este proyecto, y en una segunda sección el posible impacto del proyecto en la sociedad y la economía.
6. **Gestión del proyecto:** En este capítulo se describe la planificación y administración del proyecto, además de un análisis y estimación de costes aproximados.

2. ESTADO DEL ARTE

En este apartado se realizará un análisis exhaustivo de las tendencias y tecnologías actuales en la industria de la automoción, enfocándose en diversos aspectos claves que contribuyen a la experiencia del usuario y a la seguridad de los vehículos. Inicialmente, se explorará la experiencia del usuario en la industria de la automoción, abordando el uso de paneles táctiles, coches conectados y sistemas Heads Up Display (HUD). A continuación, se analizarán los Sistemas Avanzados de Ayuda a la Conducción (ADAS) y los vehículos autónomos, destacando las innovaciones en inteligencia artificial y aprendizaje automático, incluyendo redes neuronales. Finalmente, se revisarán trabajos relacionados y se presentarán conclusiones para establecer el contexto de la investigación y su relevancia en la mejora de la experiencia de los pasajeros.

2.1. Experiencia del Usuario en la Industria de la Automoción

La experiencia de usuario (UX) es un concepto fundamental en el ámbito del diseño y desarrollo de productos y servicios, y su importancia ha ido creciendo exponencialmente en los últimos años. En el contexto de la industria automotriz, la UX se refiere a la manera que interactúan los usuarios con los vehículos, las tecnologías que incorporan estos mismos, y cómo esas interacciones afectan a la satisfacción del usuario. [1]

Una buena experiencia de usuario no sólo implica una simple funcionalidad, ya sea ergonomía del asiento, integración de sistemas de entretenimiento, o características que mejoran la seguridad y comodidad; también implica tener en cuenta aspectos emocionales y psicológicos que influyen en la decisión de compra y la marca, como la facilidad de uso o la intuitividad de las interfaces de usuario.

Al crecer la expectativa de los usuarios para una experiencia mejor y personalizable, conlleva a los fabricantes a tomar importancia el diseño centrado en el usuario con la realización de investigaciones de mercado y pruebas de usuario para comprender sus necesidades y deseos, lo que es beneficioso para la innovación y competitividad del sector.

2.1.1. Paneles táctiles

La interacción con los componentes de un automóvil moderno es muy diferente en comparación con unas décadas atrás. En el pasado, los paneles de los coches solo contenían controles físicos sencillos y la interacción era mediante botones, diales y deslizadores para ajustar la radio o el aire acondicionado. Además, las radios solo funcionaban mediante dispositivos físicos como casetes o CDs. Posteriormente evolucionaron a pequeñas pantallas digitales que añadían una funcionalidad de GPS muy sencilla, hasta evolucionar

a lo que hay actualmente, pantallas táctiles que sustituyen e incluyen la mayoría de las funcionalidades de los botones físicos. [2]



Fig. 2.1. Evolución del panel en un Volkswagen Golf. Fuente: [1]

Esta adaptación es debido al rápido avance tecnológico de los teléfonos móviles y tablets, la comodidad en la que se interactúa pulsando o deslizando sobre las pantallas de los dispositivos, es lo que ha promovido a los fabricantes de automóviles a incorporar las mismas características de las pantallas táctiles en el lugar de los controles físicos de los coches.



Fig. 2.2. Conductor interaccionando con una pantalla táctil. Fuente: [1]

Las pantallas táctiles aportan bastantes beneficios, como pueden ser un diseño limpio y simplificado, siendo visualmente más atractivo al no contener tantos botones y diales, o flexibilidad y dinamicidad en la interfaz para añadir o mejorar funcionalidades.

Sin embargo, aunque permiten una interacción más sencilla y limpia, también presentan problemas de seguridad hacia el conductor. La distracción visual del conductor por la necesidad de desviar la vista de la carretera para interactuar con los paneles, lo que podría aumentar el riesgo de accidente o una falta de retroalimentación para las acciones que se realicen dificultaría al usuario manejar la pantalla sin mirar. Incluso en países como Reino Unido donde los vehículos tienen los volantes a la derecha, interactuar con la pantalla con la mano izquierda podría ser difícil o incómodo.

2.1.2. Coches conectados

El concepto de “coche conectado” surge de la integración de los coches en el Internet de las Cosas (IoT), al ser capaces de acceder a Internet, comunicarse con dispositivos inteligentes y recopilar datos en tiempo real. Inicialmente los automóviles contenían un número pequeño de Unidades de Control Electrónico (ECU), donde cada una era independiente de otra y realizaba una funcionalidad simple. Los coches de hoy en día pueden contener hasta 70 ECUs realizando muchas funcionalidades variadas, además de poder utilizar protocolos de comunicación que les permiten “hablar” con otros vehículos e infraestructuras, e incluso con Internet. Actualmente, incluso se ha logrado la integración de dispositivos móviles y *smartphones* en los paneles de los automóviles ofreciendo servicios avanzados de multimedia y entretenimiento. [3]

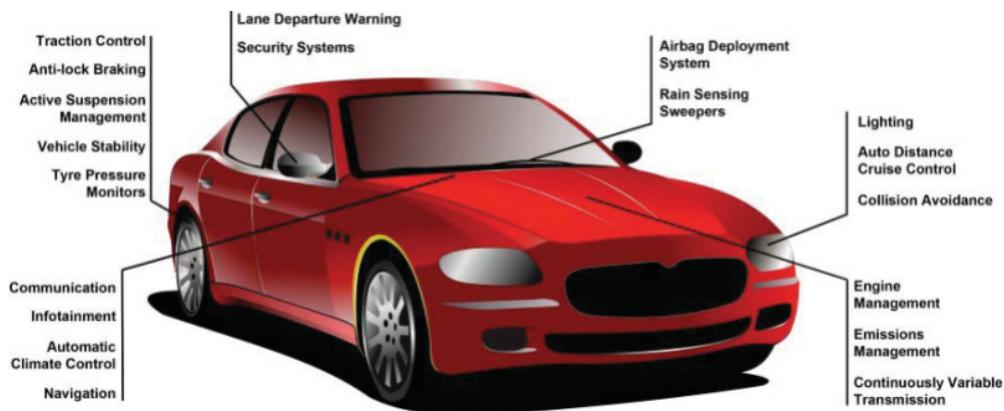


Fig. 2.3. Ejemplos de ECUs en un coche. Fuente: [3]

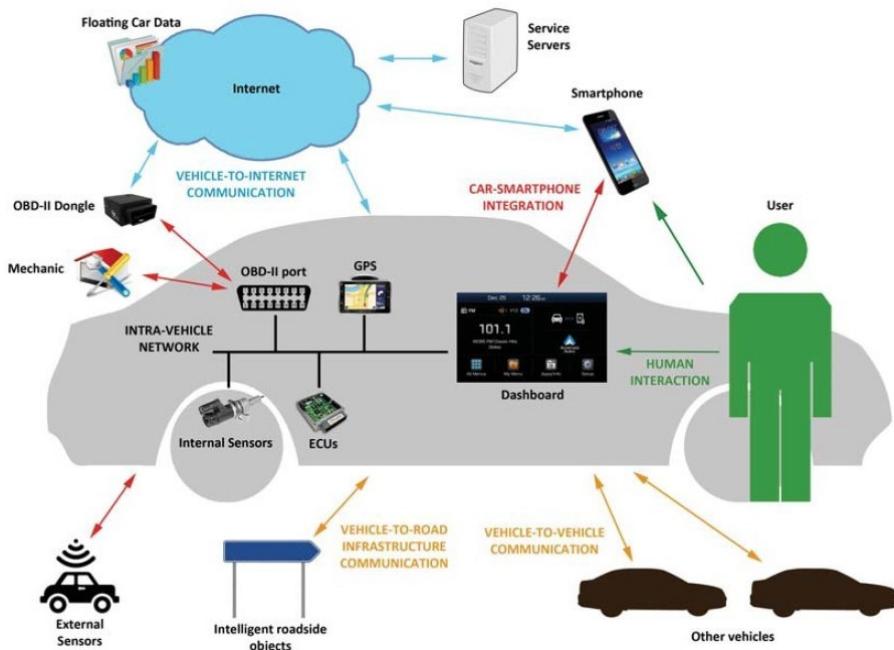


Fig. 2.4. Visión general de un sistema de coche conectado. Fuente: [3]

Al tener tantos dispositivos interconectados, las tecnologías y protocolos de comunicación en los coches conectados son fundamentales para su correcto funcionamiento, ya que permiten la interacción entre los componentes del vehículo, otros vehículos o las infraestructuras de las vías. Los diferentes tipos de comunicación que un coche conectado debe garantizar y las tecnologías disponibles para realizar dichas conexiones son las siguientes:

- **Vehicle To Sensors On Board communication (V2S) o Comunicación Intra-Vehicular:** esta comunicación se centra en la transmisión de información entre las ECUs y los sensores del vehículo, y se realiza mediante cables físicos o redes inalámbricas como Bluetooth o Ultra-Wideband. Un detalle a tener en cuenta es que aunque las redes inalámbricas ofrecen más versatilidad, presentan todavía problemas de seguridad y fiabilidad.
- **Vehicle To Vehicle Communication (V2V) o Comunicación Inter-Vehicular:** esta comunicación se centra en la transmisión de información entre diferentes coches sin la necesidad de un dispositivo remoto centralizado, utilizada principalmente para evitar accidentes, optimizar rutas, compartir información multimedia y facilitar la interacción social entre conductores. Esta tecnología de comunicación enfrenta todavía desafíos ya una constante variación en la topología de la red o la presencia de obstáculos puede dificultar su gestión o interrupciones en el flujo de los datos.
- **Vehicle To Road Infrastructure Communication (V2R):** esta comunicación se centra en la transmisión de información entre un vehículo y una infraestructura de la vía compuesta por señales de tráfico, sensores de carretera y semáforos, y utilizada principalmente para una gestión eficiente del tráfico.
- **Vehicle To Internet Communication (V2I):** esta comunicación se centra en la transmisión de información entre un vehículo e Internet, utilizada principalmente para interaccionar con sus servicios y acceder a información multimedia. Se realiza mediante infraestructuras de red celular, al igual que un dispositivo móvil, utilizando una tarjeta SIM para permitir que el vehículo se conecte a una red 3G o 4G.

En cuanto a la integración de dispositivos móviles y *smartphones* con los coches conectados, se han desarrollado soluciones software capaces de permitir a la pantalla del coche acceder a los datos del dispositivo, e incluso permite al usuario controlar aplicaciones del dispositivo en la pantalla. Esta integración ofrece varias ventajas, como pueden ser una personalización diferente para cada usuario que depende de cada dispositivo, o además facilitar a los desarrolladores de aplicaciones crear aplicaciones sobre una plataforma genérica en vez de un modelo de pantalla específico. Las soluciones de integración más extendidas son:

- **MirrorLink:** desarrollado por el Car Connectivity Consortium (CCC), y permite la conexión mediante USB o Wi-Fi Direct. Crea directrices para la interfaz gráfica de las aplicaciones, con iconos de gran tamaño facilitando una navegación rápida y sencilla.
- **Applink:** desarrollada por Ford, y permite realizar distintas acciones dentro del vehículo como mostrar mensajes en la pantalla, reproducir música y ofrecer controles por voz.
- **Apple CarPlay:** solución desarrollada para dispositivos iPhone, extendiendo las funcionalidades del sistema estándar del vehículo, y permite controlar aplicaciones mediante la pantalla, comandos de voz con SIRI, y llamadas y mensajes con manos libres.
- **Android Auto:** solución desarrollada por la Open Automotive Alliance, y centrada en la conexión con dispositivos Android. Proporciona un tablero simple que permite el acceso a sistemas de navegación, como por ejemplo Google Maps, además de música, llamadas telefónicas, comandos de voz, etc.

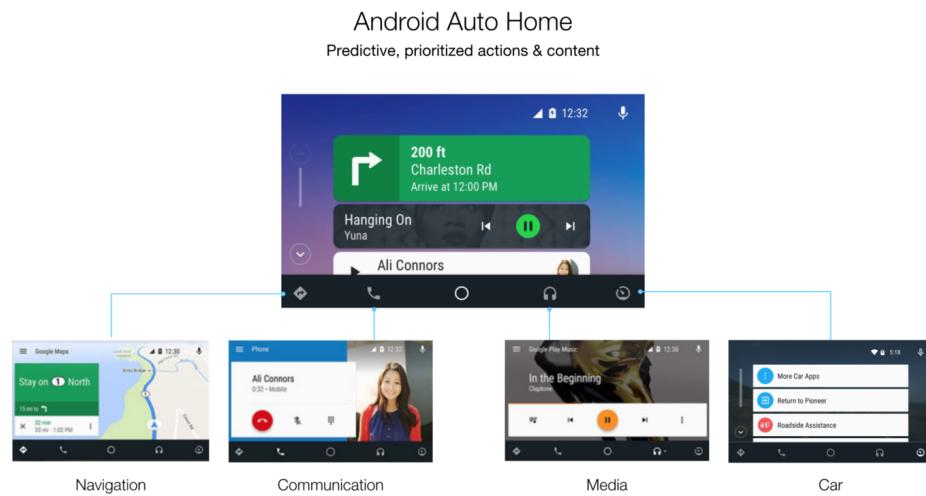


Fig. 2.5. Ejemplos de interfaces de Android Auto en la pantalla del coche.

Actualmente, de las tecnologías anteriores las más populares y mayoritariamente implementadas son las dos últimas, Apple CarPlay y Android Auto, debido a que tienen un mayor número de funcionalidades y tienen un mayor número de dispositivos con dicho sistema operativo compatible.

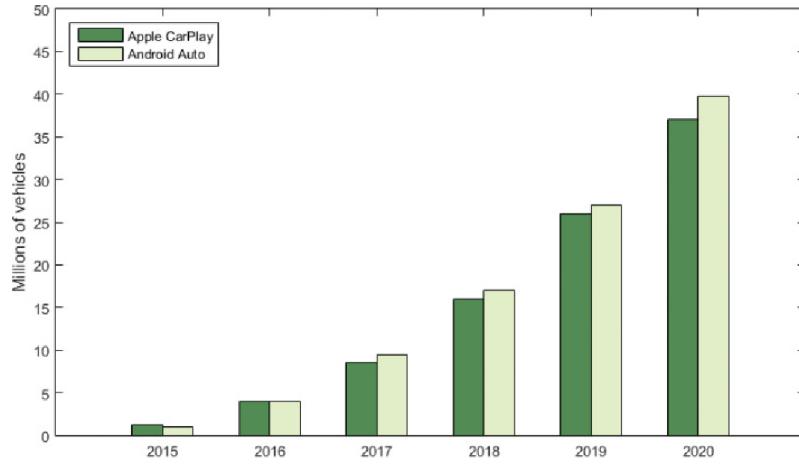


Fig. 2.6. Ventas de coches con Apple CarPlay vs Android Auto. Fuente: [3]

2.1.3. Heads Up Display (HUD)

Los sistemas Head-Up Display (HUD) fueron desarrollados originalmente para aviadores durante la Primera Guerra Mundial, los cuales servían para percibir aviones enemigos, y que posteriormente fueron evolucionando para mostrar más información como altitud, velocidad del viento, indicador de horizonte, etc. Actualmente se han introducido también en vehículos a tierra, los coches, mostrando datos esenciales sobre el parabrisas del vehículo. [4]



Fig. 2.7. Vista de un HUD de un Jaguar. Fuente: [4]

La estructura de un HUD se compone de tres componentes: el combinador, la superficie donde se proyecta la imagen; el proyector, el componente que genera y dirige la imagen hacia el combinador; y la unidad de procesamiento, que gestiona los datos obtenidos y controla la protección.

La información que es proporcionada por un HUD de un automóvil está centrada en información relacionada con la conducción y el estado del vehículo: características de la carretera como límites de velocidad y condiciones del camino, o navegación con mapas

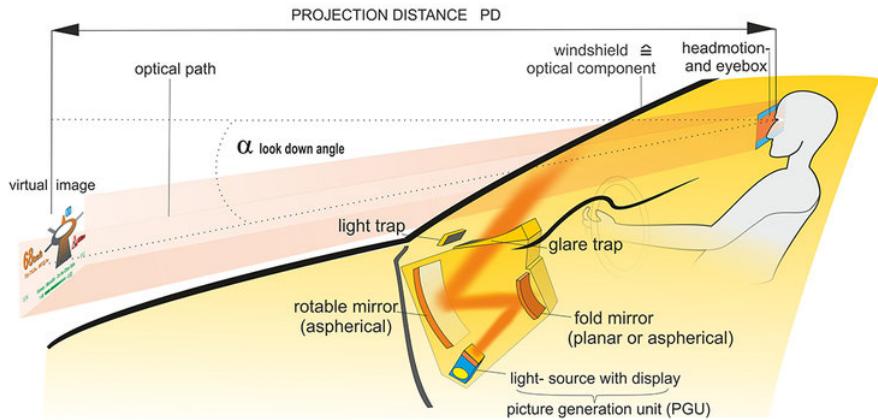


Fig. 2.8. Funcionamiento y componentes de un HUD. Fuente: [4]

y direcciones. Todo esto es útil para mantener al conductor informado, y proporcionar advertencias en caso de eventos peligrosos que puedan incurrir a la seguridad.

También se han desarrollado dispositivos HUD comerciales, que pueden instalarse de manera externa, aunque tienen limitaciones para usar la información de navegación al no ser integrados.

Algunos vehículos modernos incluyen incluso funcionalidades como visión nocturna, mediante el uso de cámaras infrarrojas para mejorar la visibilidad en condiciones de poca luz, y proyectando esa información en el HUD.



Fig. 2.9. Visión de cámara de infrarrojos en el HUD. Fuente: [4]

También existen casos experimentales donde se implementan en los espejos retrovisores para proporcionar información de vistas traseras.



Fig. 2.10. Visión de HUD en espejos retrovisores. Fuente: [4]

2.2. Sistemas Avanzados de Ayuda a la Conducción (ADAS)

Los Sistemas Avanzados de Asistencia al Conductor son un conjunto de tecnologías, combinando hardware y software, que ayudan a tener una conducción segura y eficiente. El objetivo principal es reducir el riesgo de accidentes y mejorar la experiencia de conducción mediante el uso de sensores, cámaras, radares y distintos dispositivos que monitorean el entorno alrededor del vehículo y avisan de manera activa al conductor. De esta manera, no sólo mejoran la seguridad de los ocupantes de un vehículo, también proporcionan más seguridad a otros ocupantes de las vías como pueden ser los peatones o ciclistas. [5]

Debido a la importancia y eficacia de estos sistemas para la seguridad, la Unión Europea aprobó en 2019 el Nuevo Reglamento de Seguridad 2019/2144 [5] que recogía una serie de requisitos a cumplir para garantizar la seguridad de los vehículos y las personas. En España se implantó el 6 de julio de 2022 dicho Reglamento, y obliga a todos los coches nuevos de fábrica a incorporar al menos los siguientes ocho sistemas ADAS:

- **Detector de Fatiga y Somnolencia (DDR):** este sistema de ADAS se enfoca a prevenir distracciones del conductor debido a somnolencia, cansancio o falta de atención. Utiliza distintos tipos de tecnologías como: sensores en el volante para analizar la presión que se ejerce sobre él o la brusquedad de las correcciones de dirección; una cámara de reconocimiento facial que detecta el exceso de parpadeo o la visión de la mirada; o temporizador en el motor que recoge el tiempo que lleva en marcha. Si el sistema detecta anomalías o un tiempo de marcha largo, emite una señal o texto por los paneles del coche para avisar al conductor.

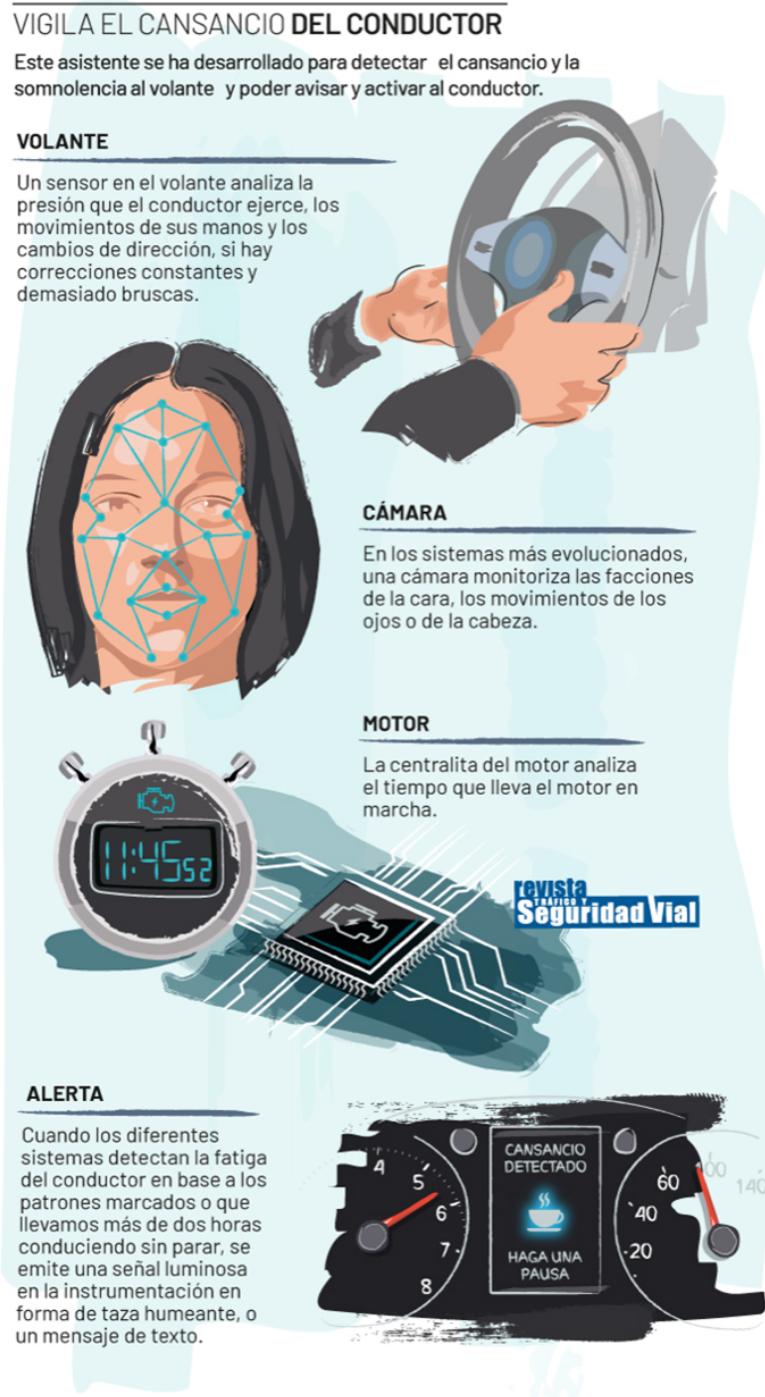


Fig. 2.11. Funcionamiento del sistema DDR. Fuente: [6]

- **Asistente de velocidad inteligente (ISA):** este sistema ADAS se enfoca en limitar la velocidad máxima a la que se circula mediante una cámara interior que lee las señales de las vías. La velocidad permitida es mostrada en el cuadro de mando del conductor, y el software actúa sobre el motor o el acelerador impidiendo que supere esa velocidad máxima, aunque se puede desconectar mediante un botón o pisando con fuerza el acelerador.



Fig. 2.12. Funcionamiento del sistema ISA. Fuente: [6]

- **Cámara Trasera con Detección de Tráfico Cruzado (RCTA):** este sistema ADAS se enfoca en facilitar la visión del conductor en la marcha atrás con cámaras que muestran la imagen en el panel. También detecta posibles obstáculos que puedan estar en la trayectoria de la maniobra, advierte al conductor en la pantalla o en el retrovisor, y si el conductor no reacciona puede llegar a detener el coche de manera automática.



Fig. 2.13. Funcionamiento del sistema RCTA. Fuente: [6]

- **Caja negra (EDR):** este sistema ADAS se enfoca en recopilar información, en caso de accidente, de los momentos anteriores y posteriores a él, para poder evaluar las posibles causas y estado de los sistemas del vehículo. El EDR está ubicado debajo del asiento del conductor, del tamaño aproximado de un teléfono móvil, y se puede conectar a un ordenador para obtener la información necesaria.



Fig. 2.14. Funcionamiento del sistema EDR. Fuente: [6]

- **Alerta de cambio involuntario de carril (LDW):** este sistema ADAS se enfoca en controlar que el vehículo no se desvía del carril por el que circula por una posible falta de atención del conductor. Utiliza una cámara frontal que detecta de forma continua las líneas del carril por el que circula, y en caso de que no se haya activado ningún intermitente y el vehículo esté saliendo del carril, avisa mediante una señal en el cuadro, un sonido o vibraciones. Además, si dispone de algún sistema de mantenimiento de carril, podría activarse para mantener el vehículo correctamente en el carril.



Fig. 2.15. Funcionamiento del sistema LDW. Fuente: [6]

- **Sistema de frenada de emergencia (BAS y EBA):** este sistema ADAS se enfoca en reducir el riesgo de colisión por alcance. Al realizar una frenada brusca, detecta si la intensidad con la que se pisa es la adecuada, y si no es suficiente el sistema interacciona frenando a la máxima intensidad posible. Además dispone de distintos sensores que permite distinguir a las personas y realiza un aviso, si el conductor no reacciona al aviso se activa el freno de manera automática.

ACTIVA AUTOMÁTICAMENTE EL FRENO

Reducir el riesgo de colisión por alcance o en curva y mitigar sus consecuencias, es el objetivo de este sistema que resulta un excelente aliado del conductor



Fig. 2.16. Funcionamiento del sistema BAS y EBA. Fuente: [6]

- **Bloqueo de vehículo con alcoholímetro (Alcolock):** este sistema ADAS se enfoca

en el control de la conducción tras la ingesta de alcohol. El vehículo dispone de un sistema de alcoholímetro, conectado al motor de arranque, y que impedirá al conductor arrancar el motor si la tasa de alcohol analizada por la boquilla sea superior al índice permitido. El sistema tendrá implementado sensores con tecnologías de reconocimiento de huellas dactilares, detección facial o fotografías digitales, que serán guardadas en memoria al igual que el historial de análisis, que servirá para identificar al usuario e impedir la falsificación o el fraude.

CON ALCOHOL, NO ARRANCA

Todos los coches a partir de 2022 dispondrán de preinstalación para el alcoholímetro con bloqueo de la puesta en marcha. Conocido como alcolock



Fig. 2.17. Funcionamiento del sistema Alcolock. Fuente: [6]

- **Alerta de cinturón en las plazas traseras:** este sistema ADAS se enfoca en controlar y avisar el uso correcto del cinturón de seguridad en las plazas traseras. El sistema determina mediante un sensor de carga si la plaza está ocupada, y mediante un sensor de la hebilla si está bien cerrado. En caso de que el vehículo recorra una determinada distancia, un determinado tiempo o supere los 25 km/h y no se haya puesto el cinturón de una manera correcta, el sistema avisa automáticamente mediante señales visuales y acústicas.



Fig. 2.18. Funcionamiento del sistema de Alerta de cinturón en las plazas traseras. Fuente: [6]

2.3. Vehículos Autónomos

El gran avance de la tecnología, en el ámbito del IoT y los sistemas embebidos, hacen que los vehículos se vuelvan cada vez más inteligentes, hasta llegar al punto de poder realizar acciones de forma autónoma sin intervención de las personas. Los vehículos autónomos (AV) tienen como objetivo mejorar la seguridad en las vías, facilitar las maniobras diarias de los conductores y disminuir el riesgo de accidentes debido a cansancio o fatiga, y optimizar recursos y tiempo de manera eficiente. [7]

Un vehículo autónomo puede clasificarse en diferentes niveles de autonomía, dependiendo del grado o acciones que llega a realizar el vehículo de forma autónoma en comparación con el conductor. Existe una normativa estandarizada, definida por la Sociedad de Ingenieros Automotrices (SAE), el estándar J3016 o Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems [8], en el que se describe seis niveles de autonomía [9]:



SAE J3016™ LEVELS OF DRIVING AUTOMATION

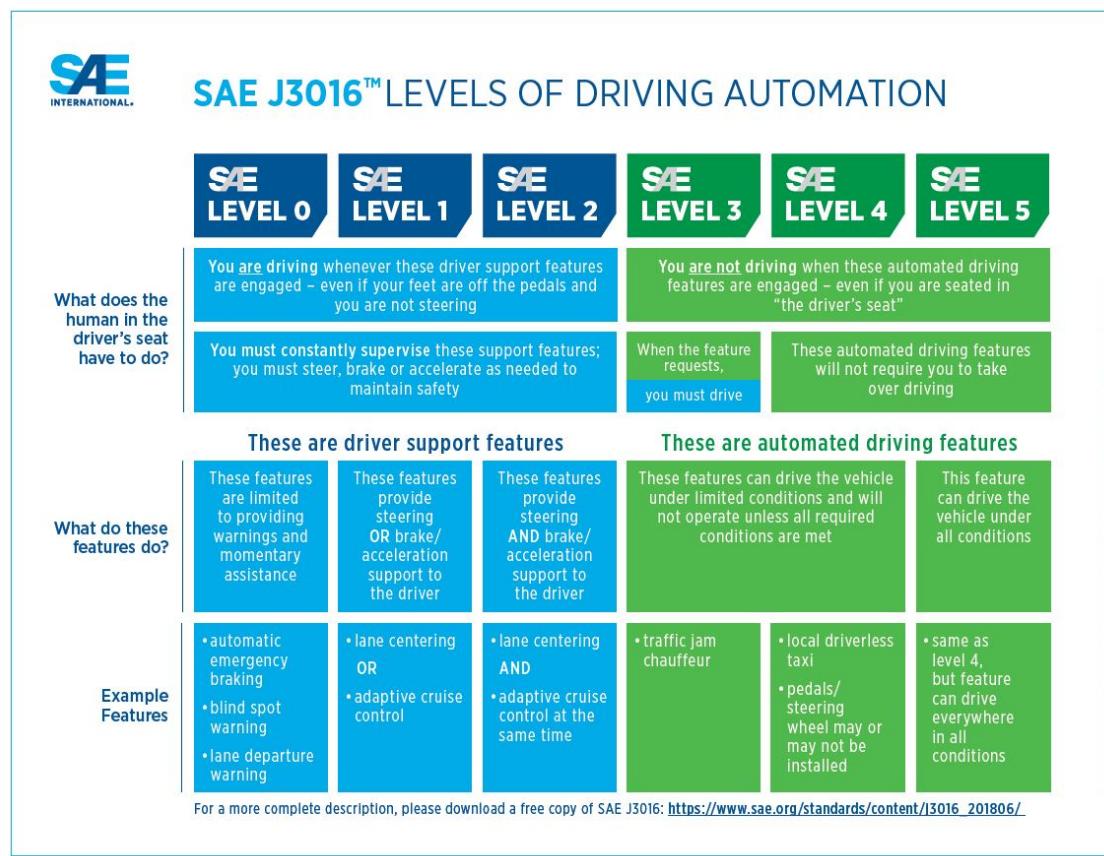


Fig. 2.19. Niveles de autonomía según el estándar SAE J3016. Fuente: [9]

- **SAE Nivel 0 (Sin Automatización):** en este nivel, el conductor es responsable de todas las acciones relacionadas a la conducción, como son la aceleración, el frenado, el manejo de dirección y el uso de señales. El vehículo puede tener instalado sistemas de advertencia, como alertas de colisión o de objetos en punto ciego, pero en ningún caso realiza alguna acción controlando el vehículo por el conductor.
- **SAE Nivel 1 (Asistencia al Conductor):** en este nivel, se dispone de sistemas que pueden controlar funciones específicas del vehículo, como la aceleración (control de crucero adaptativo), el frenado (frenado de emergencia) o el manejo de dirección (sistema de mantenimiento de carril). El conductor debe estar atento y en supervisión constante para tomar el control de vuelta de los pedales en caso de necesidad.
- **SAE Nivel 2 (Automatización Parcial):** en este nivel, el vehículo puede además controlar la dirección de forma autónoma y simultánea con la aceleración y el frenado, pero el conductor debe estar al igual en supervisión constante sobre el entorno y tomar el control en caso necesario.
- **SAE Nivel 3 (Automatización Condicional):** en este nivel, el vehículo es capaz de percibir el estado del entorno de manera autónoma, y controlar todas las funciones de conducción en determinadas condiciones simples, como en autopistas. Aun

así, en este caso el conductor debe estar en supervisión para tomar el control si el sistema lo solicita.

- **SAE Nivel 4 (Alta Automatización):** en este nivel, el vehículo puede encargarse de todas las funciones de forma autónoma en situaciones específicas y bien definidas, como por ejemplo en autopistas o áreas urbanas. El conductor sigue necesitando tener presencia en determinadas circunstancias, pero en este nivel ya existe la posibilidad de no utilizar volante ni pedales.
- **SAE Nivel 5 (Automatización Completa):** en este nivel, el vehículo es capaz de encargarse de todas las funciones en cualquier situación sin la necesidad de presencia humana.

Los vehículos autónomos están equipados con una gran cantidad de sensores y sistemas que les permiten lograr la percepción del entorno y poder conducir de una manera segura [9]:

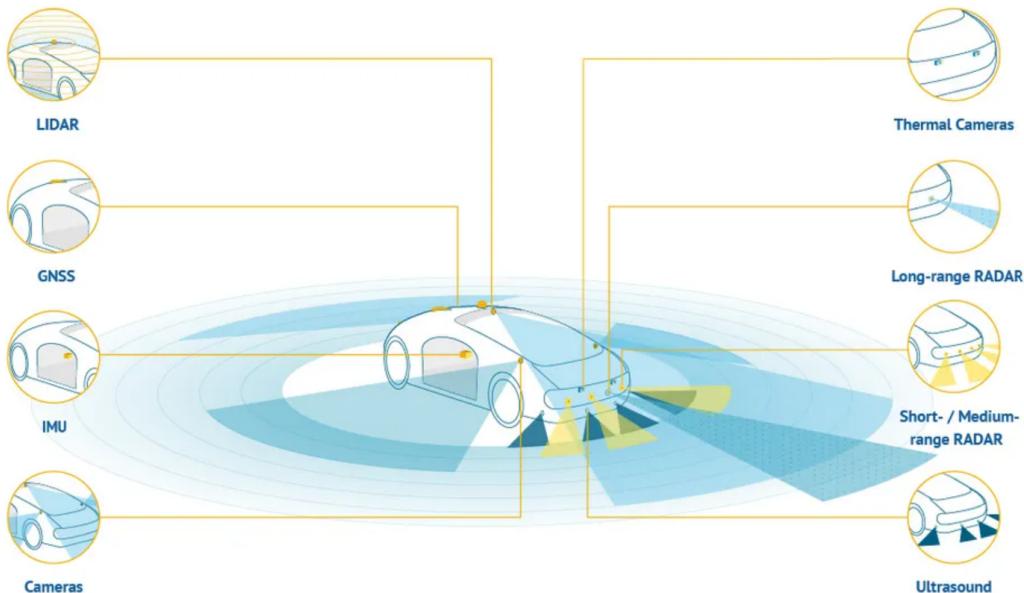


Fig. 2.20. Sensores en un vehículo autónomo. Fuente: [9]

- **Light Detection and Ranging (LIDAR):** sensor que funciona mediante detección láser, analizando la luz reflejada que se envía. Este sensor es capaz de detectar los objetos de alrededor del vehículo y crear mapas virtuales de 2D o 3D.
- **Radio Detection and Ranging (RADAR):** sensor que funciona mediante ondas electromagnéticas que detectan los obstáculos o vehículos cercanos. Pueden ser de largo alcance, medio alcance o corto alcance, y se usan en sistemas como el control de crucero adaptativo, la advertencia de colisión o la detección del ángulo muerto.
- **Global Navigation Satellite System (GNSS):** conjunto de sistemas que funcionan mediante señal de los satélites, como por ejemplo el GPS, y proporciona información precisa sobre la ubicación y la hora.

- **Sensores Ultrasónicos:** sensores que detectan objetos en una distancia muy cercana al vehículo, utilizados en sistemas de asistencia al estacionamiento.
- **Unidad de Medición Inercial (IMU):** sensor que mide el movimiento lineal y angular del vehículo utilizando acelerómetros y giroscopios. La información que recibe se combina con los datos del sistema GPS para obtener una mayor precisión en la posición.
- **Cámaras térmicas:** cámaras que detectan el calor, con un rango más amplio que los faros, y son útiles para la detección de obstáculos, ya sean personas, animales u otros coches, en condiciones difícilmente perceptibles, como por la noche, en zona de humo o con un brillo solar intenso.
- **Cámaras de video:** cámaras de alta resolución, que monitorean y transmiten de manera constante imágenes de la vía, usadas para procesar información relevante como señales de tráfico, peatones o ciclistas.

El conjunto de sensores, junto a los sistemas ADAS, son lo que permiten a los conductores circular de manera más segura, y lo que hace posible que los vehículos autónomos puedan realizar tareas de conducción sin la intervención del conductor.

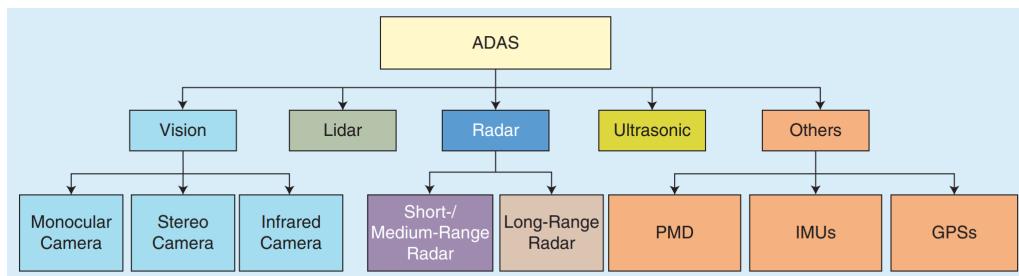


Fig. 2.21. Clasificación de sensores usados en sistemas ADAS. Fuente: [10]

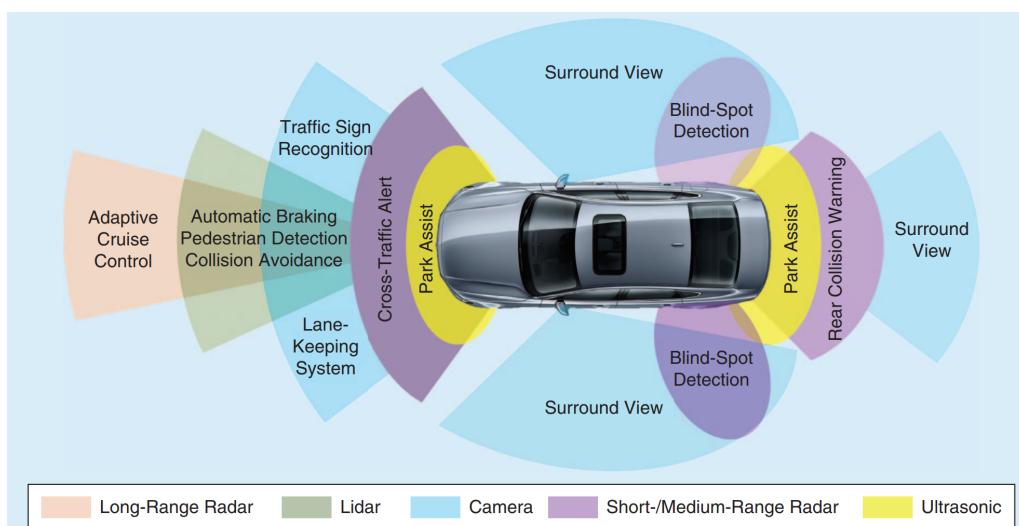


Fig. 2.22. Rangos sobre los que actúan los sensores. Fuente: [10]

2.4. Inteligencia Artificial y Aprendizaje Automático

La Inteligencia Artificial (IA) se define como “capacidad de las máquinas para imitar la inteligencia de las personas humanas en la realización de tareas” [11]. La investigación en IA está centrada en el desarrollo de dispositivos que perciben su entorno, resuelven problemas, y aprenden de los errores que sirvan posteriormente, para alcanzar un objetivo específico.

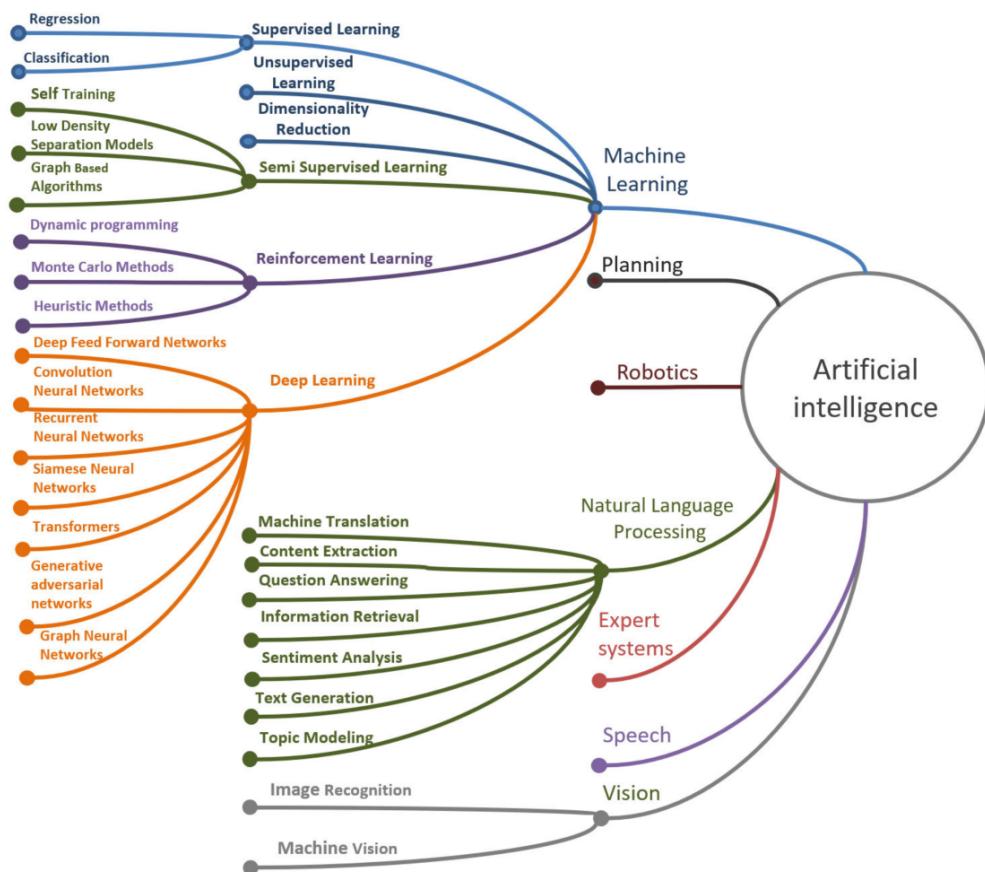


Fig. 2.23. Campos de la Inteligencia Artificial. Fuente: [12]

El Aprendizaje Automático (ML), por otra parte, es un campo de la IA que está centrado en el desarrollo de esos algoritmos que les permiten aprender y hacer predicciones a partir de los datos o problemas. Es definido como “capacidad de las máquinas para aprender sin ser programadas específicamente” [11]. En la práctica, se puede definir como la construcción de modelos matemáticos mediante unos datos de muestra, y realizar con esos modelos predicciones o tomar decisiones sobre datos nuevos. [11]

Los métodos de ML más comunes que se utilizan para aprender y realizar predicciones son los siguientes [11] [12]:

- **Aprendizaje Supervisado (SL):** método en el que se utilizan datos etiquetados para el entrenamiento del modelo. Esto quiere decir que para cada dato de entrada existe una salida final que es conocida, y el objetivo del modelo es aprender a

unir cada entrada con su salida. Tras tener el modelo entrenado, este será capaz de predecir las salidas no conocidas de nuevas entradas.

- **Aprendizaje No Supervisado (UL):** método en el que se utilizan datos no etiquetados para el entrenamiento del modelo. En este caso, el modelo pretende encontrar patrones ocultos o estructuras en los datos de entrada sin conocer la salida.
- **Aprendizaje Semi-Supervisado (SSL):** método en el que se utiliza una combinación de datos etiquetados y no etiquetados para el entrenamiento del modelo. Este método es útil en caso de que la obtención de los datos etiquetados sea costosa.
- **Aprendizaje por refuerzo (RL):** método en el que se entrena un agente, que interactúa con el entorno realizando determinadas acciones y obteniendo recompensas o penalizaciones dependiendo del resultado de la acción, y aprende con el objetivo de maximizar la recompensa acumulada. Es útil en campos como la robótica, los juegos o la navegación autónoma.
- **Aprendizaje Profundo (DL):** conjunto de métodos más complejos que utilizan redes de neuronas artificiales con múltiples capas ocultas para aprender patrones en un conjunto de datos muy grande. Las arquitecturas de DL más conocidas son redes neuronales recurrentes (RNN) y redes neuronales convolucionales (CNN).

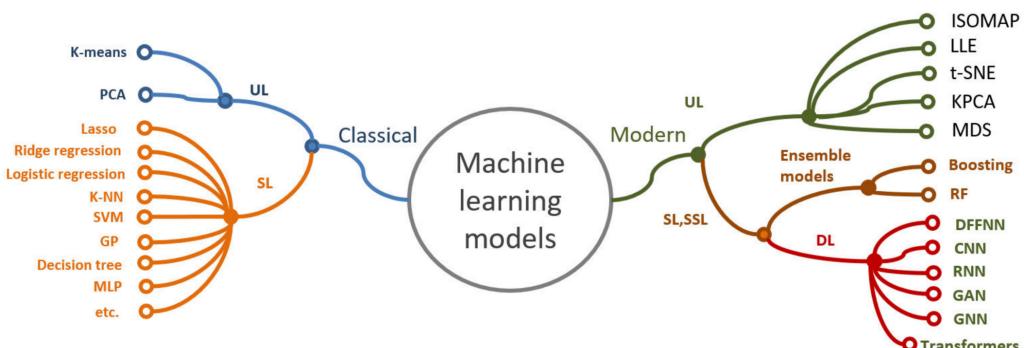


Fig. 2.24. Modelos de Aprendizaje Automático. Fuente: [12]

2.4.1. Redes Neuronales

Una red de neuronas artificial (ANN), o simplemente red de neuronas, es un sistema basado en múltiples nodos llamados neuronas, distribuidas en varias capas. Las capas principales de una red de neuronas son la capa de entrada que recibe los datos, las capas ocultas que procesan los datos, y la capa de salida que devuelve los resultados. [13]

Cada neurona de una capa recibe entradas, procesa esas entradas y devuelve una salida que es enviada a la capa siguiente. El proceso de un dato en una neurona se realiza a través de una función de activación, función que introduce características no lineales en la red para que pueda aprender relaciones complejas entre el dato de entrada y salida.

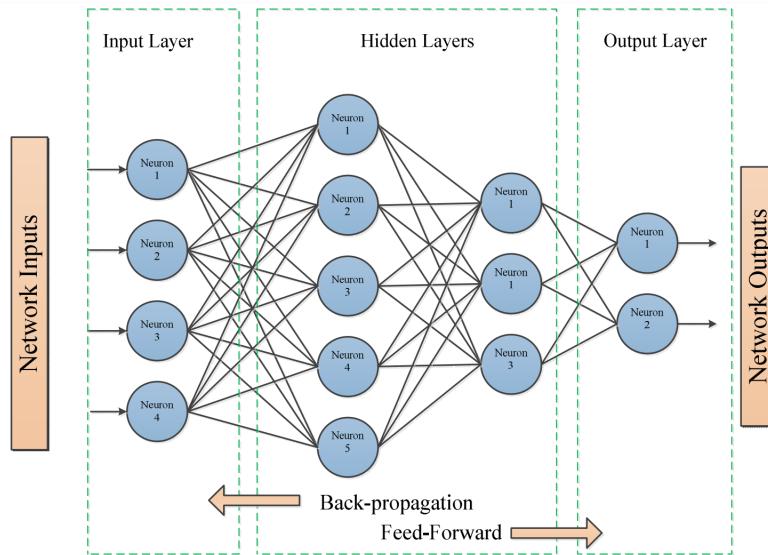


Fig. 2.25. Artificial Neural Network basic architecture. Fuente: [13]

Redes Neuronales Recurrentes (RNN)

Las redes de neuronas recurrentes (RNN) son un tipo de red neuronal que destacan por tener conexiones en forma de ciclos en la red, permitiéndoles mantener una “memoria” de los estados anteriores, por lo que son eficaces con el tratamiento de tareas donde tiene importancia el contexto temporal. [12] [13]

Las RNN procesan los datos de manera secuencial: en cada estado de procesamiento, se toma un valor de entrada y su estado anterior, y a partir de ambas genera una salida y actualiza su estado. Este modo de procesamiento en forma de ciclo es lo que permite a la red mantener la “memoria” de la información a lo largo de todo el proceso, y lo que es relevante para tareas como el procesamiento del lenguaje natural (NLP), donde la comprensión de una palabra depende del contexto anterior.

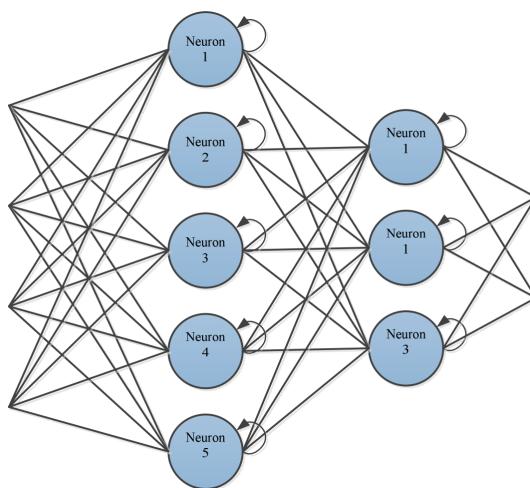


Fig. 2.26. Recurrent Neural Network basic architecture. Fuente: [13]

Redes Neuronales Convolucionales (CNN)

Las redes de neuronas convolucionales (CNN) son un tipo de red neuronal que destacan por poder procesar y analizar datos que tengan una estructura de cuadrícula, como por ejemplo las imágenes. [12] [13]

Una CNN está compuesta por varias capas de distintos tipos que transforman la entrada recibida de una manera específica:

- **Capa convolucional:** capa que aplica filtros o núcleos a los datos de entrada, siendo el filtro una matriz pequeña que se desliza sobre el dato, y realiza una convolución sobre los datos para obtener un mapa de características. Estas características son las que permiten a la red detectar los patrones en la imagen, como bordes o texturas.
- **Capa de pooling o subsampling:** capa que aplica una reducción de dimensionalidad sobre el mapa de características, generalmente siendo una operación de max-pooling tomando el valor máximo de cada región. Esta capa sirve para reducir la cantidad de parámetros y controlar el sobreajuste.
- **Capas completamente conectadas:** capas que se sitúan al final de la red, y sirven para procesar las características obtenidas de las capas anteriores y realizar una clasificación final. Cada neurona de esta capa se conecta a todas las neuronas de la capa anterior.

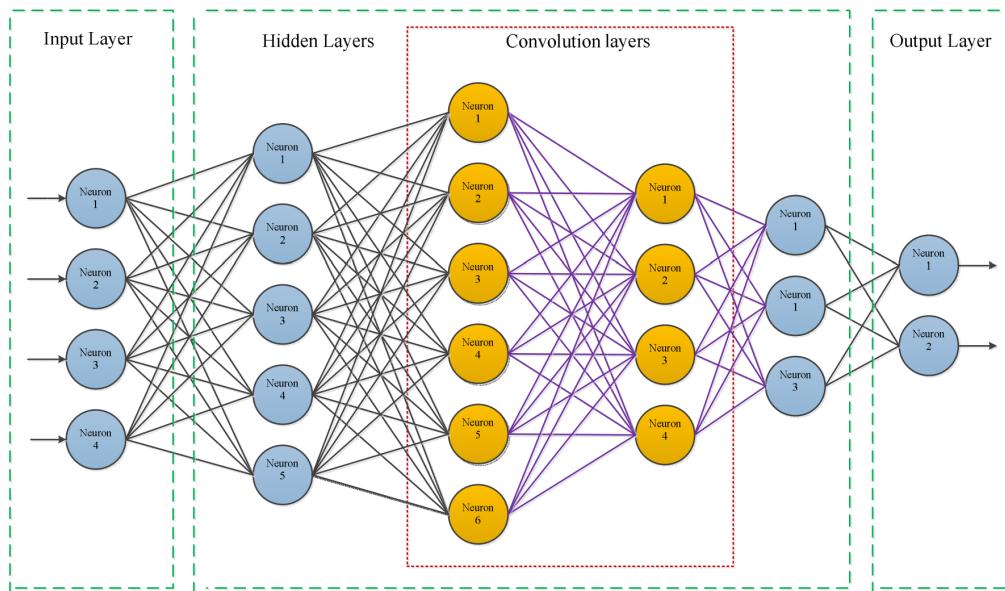


Fig. 2.27. Convolutional Neural Network basic architecture. Fuente: [13]

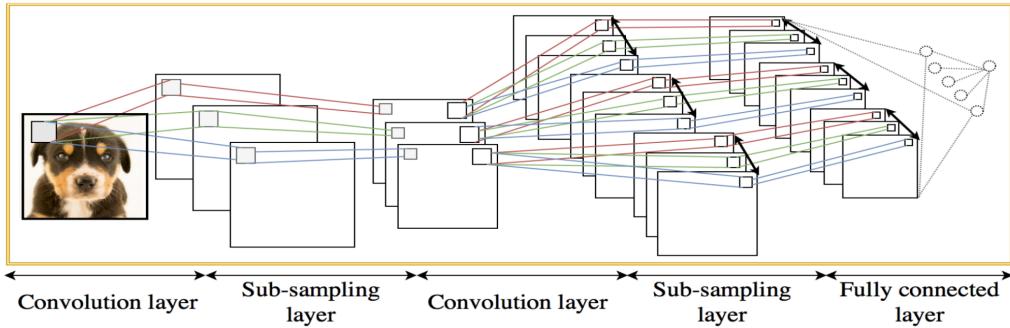


Fig. 2.28. Convolutional Neural Network layers visualization. Fuente: [14]

2.5. Trabajos Relacionados

El reconocimiento de emociones ya sea de los ocupantes o del conductor en un vehículo es un área de desarrollo fundamental para la industria automotriz. Una capacidad para identificar y responder eficaz puede mejorar considerablemente la seguridad o la experiencia de usuario de los ocupantes del vehículo. Al ser un área de investigación que ha ido avanzado a lo largo de los años, se han realizado muchos estudios destacados, llegando al estado actual con sistemas tecnológicos cada vez más sofisticados.

Un artículo de investigación interesante y reciente es “Drivers’ Comprehensive Emotion Recognition Based on HAM” [15] realizado en 2023, basado en mejorar la precisión de modelos de aprendizaje profundo para reconocer emociones, en este caso del conductor, utilizando señales audiovisuales. La metodología utilizada es una combinación híbrida de diferentes redes neuronales, y un entrenamiento y validación con distintos conjuntos de datos, y lograron una precisión del 79.60 % en el conjunto de datos FER2013 superando a otros modelos ya existentes.

Otro artículo destacado es “AI-Powered In-Vehicle Passenger Monitoring Using Low-Cost mm-Wave Radar” [16] publicado en 2021, en el que se desarrolla un sistema de detección y monitoreo de los pasajeros de un vehículo. Concretamente, trata de identificar los asientos ocupados y contar el número de pasajeros que hay mediante un radar de onda milimétrica generando mapas de calor, y procesando la información con algoritmos de aprendizaje automático. En comparación con los sistemas tradicionales, este sistema permite detectar pasajeros con más precisión, como por ejemplo al detectar múltiples pasajeros en asientos adyacentes, además de ser una solución económica con el uso de radares de bajo costo.

En cuanto a artículos en relación a la experiencia de usuario, un artículo destacado es “Mobility Prediction-Based Optimisation and Encryption of Passenger Traffic-Flows Using Machine Learning” [17] publicado en 2020. Este artículo se centra en desarrollar un modelo predictivo de aprendizaje automático para mejorar el rendimiento de infraestructuras ferroviarias, en este caso del metro de Londres, abordando problemas como la congestión, los accidentes, la sobrecarga de vagones o el mantenimiento. Utilizan senso-

res de identificación por radiofrecuencia (RFID) que registran los datos de pasajeros en tiempo real mediante tecnología 5G, y con un modelo de redes neuronales artificiales predicen los patrones de flujo de los pasajeros. Además todos los datos enviados en tiempo real son encriptados lo que asegura su seguridad y privacidad. Realizaron simulaciones con datos reales del metro de Londres y el modelo alcanzó una precisión del 91.17 % en la predicción del flujo de movilidad de los pasajeros.

Otro ejemplo de artículo es “Estimating passenger experience from vehicle data: Pre-conditions for using machine learning” [18] publicado en 2020. Este artículo se centra en evaluar métodos de aprendizaje automático para estimar de forma automática la experiencia de los pasajeros en los autobuses eléctricos, utilizando datos obtenidos con sensores en el autobús y encuestas a los pasajeros sobre la experiencia del viaje. Las implicaciones de este estudio puede ser interesante para los fabricantes de vehículos, ya que podrían mejorar su diseño y desarrollo, o para los operadores de los transportes públicos, ya que podrían mejorar la formación de conductores o la planificación de rutas en función de la satisfacción del pasajero.

Otro artículo, también relacionado a mejorar la experiencia de usuario sobre el uso de autobuses como medio de transporte público, es “A Real-Time Machine Learning-Based Public Transport Bus-Passenger Information System” [19] publicado en 2021. El objetivo de este estudio es desarrollar un sistema de información en tiempo real que muestra la ubicación de los autobuses para mejorar su puntualidad y confiabilidad, y optimizar la asignación de recursos mediante técnicas de machine learning, usando datos diarios para predecir demandas de pasajeros en diferentes rutas. Esto servirá para estimar de una manera más exacta el tiempo de llegada de un autobús y por lo tanto optimizar tiempo, además de asignar de una manera más eficiente autobuses a las diferentes rutas que lo necesiten.

2.6. Conclusiones del Estado del Arte

Tras haber realizado un análisis e investigación del Estado del Arte y los artículos realizados en relación al tema, se ha podido observar que en la evolución del automóvil el factor de la seguridad es un papel muy importante e innovador, sobre todo en las tecnologías ADAS. Esto es debido sobre todo a los avances tecnológicos de los sensores, las mejoras en las capacidades de cómputo y el desarrollo de algoritmos más avanzados de IA, todo esto de relación muy importante si en un futuro se quiere transicionar de vehículos tradicionales a vehículos autónomos.

La mayoría de los sistemas ADAS actualmente están diseñados para asistir al conductor, mientras que no hay muchos que sean enfocados a los pasajeros. Pero la experiencia del usuario, ya sea centrada en la seguridad como en la comodidad, debería tener igual importancia, y más en los vehículos autónomos donde la presencia del conductor será cada vez menor.

Tras obtener estas conclusiones del análisis, con este proyecto se pretende proponer un sistema, cuyo diseño se centra en mejorar la experiencia de usuario de los ocupantes de un vehículo, y que funcione a partir de análisis en tiempo real de sus emociones y reaccionar de una manera u otra interactuando con las funcionalidades de los vehículos, como son el sonido de la música o radio, el aire acondicionado, la comodidad de los asientos, o integrarse en sistemas ADAS que sirvan para mejorar la experiencia de los ocupantes.

3. DISEÑO DEL SISTEMA

Al igual mencionado anteriormente, los avances hoy en día en la industria automotriz sobre los sistemas de seguridad y experiencia al usuario están mayoritariamente enfocados al conductor, y las investigaciones y proyectos en los que se tienen en cuenta a los acompañantes son relacionados al sector de los transportes públicos.

Por el anterior motivo, se propone un sistema cuyo diseño está enfocado para los ocupantes de un vehículo privado. Concretamente, el uso principal del sistema es mejorar la experiencia de un ocupante mediante el control automático de funcionalidades de multimedia de un vehículo dependiendo del estado emocional en el que se encuentren los ocupantes.

El sistema también estaría abierto a otras aplicaciones en sistemas ADAS, como por ejemplo restringir la velocidad máxima que puede alcanzar un vehículo en caso de que el estado emocional de un acompañante sea de enfado o estrés, ya que podría influir en el comportamiento del conductor, y los sistemas actuales solo están implementados para detectar el estado del conductor y no de los acompañantes. Pero esto es solo una posible idea de aplicación posible, que no se abordará en este proyecto.

Por lo tanto, la arquitectura de nuestro sistema estará compuesta de diferentes módulos clave, cada uno con una funcionalidad separada, y que juntos forman el sistema final que permitirá el reconocimiento de emociones y la adaptación del resultado sobre las funcionalidades del vehículo:

1. Módulo de obtención de imágenes en tiempo real

- Las imágenes se obtendrán mediante cámaras de alta resolución instaladas en los coches para capturar imágenes de las caras y emociones de los acompañantes. El software para su funcionamiento se encargará de la frecuencia de captura, la calidad de captura, la cantidad de imágenes, y el pre-procesamiento y envío en tiempo real de las imágenes, ya sea por el bus CAN del vehículo o por conexión inalámbrica.

2. Módulo de predicción y detección de emociones

- Las imágenes recibidas serán analizadas por un modelo de red neuronal convolucional (CNN) pre-entrenado que detectará las emociones de esas imágenes, y el resultado será enviado a la API o interfaz de software encargada de la interacción con el vehículo.

3. Módulo de interacción

- El resultado obtenido de los análisis de las imágenes será utilizado para decidir las acciones a realizar, y se comunicará con el sistema de interacción, que

controlará las funcionalidades del vehículo según el resultado: por ejemplo si los acompañantes están alegres subir el volumen de la música y activas las luces de ambiente, si están tristes cambiar a una música más animada, si están enfadados bajar el volumen de la música, si están cansados o durmiendo bajar la música y el aire acondicionado, etc. Dependiendo de cada una de las posibilidades que prediga en modelo, podrían haber diferentes combinaciones, incluso podría ser personalizable por el mismo usuario.

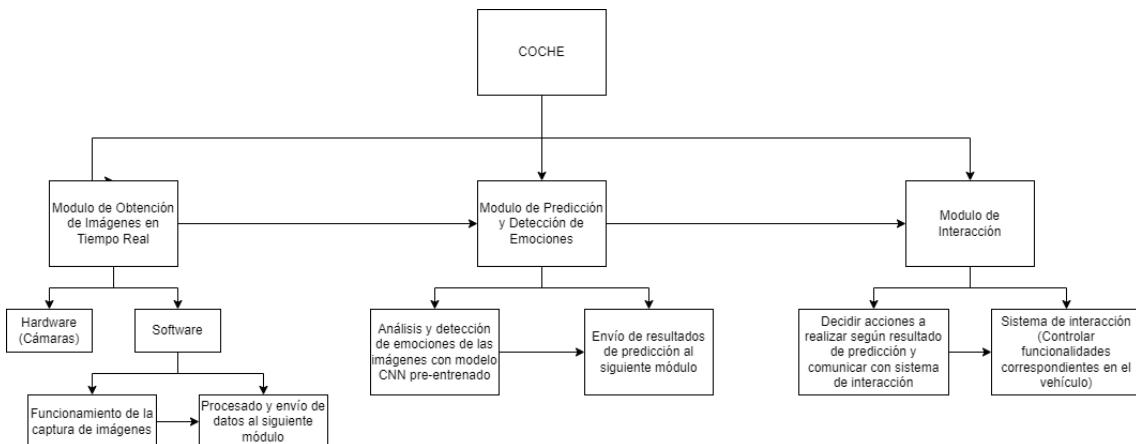


Fig. 3.1. Esquema de relación y funcionalidad de los módulos.

3.1. Análisis de requisitos

El sistema deberá cumplir una serie de requisitos funcionales, que definen las funcionalidades y características que el sistema debe tener para alcanzar los objetivos del proyecto, y no funcionales que definen el comportamiento del sistema enfocado a la calidad de servicio y funcionamiento.

Los requisitos se definen en formato de tabla, siendo cada requisito identificado por RF si es funcional y RNF si es no funcional, seguido de un número de requisitos. Cada definición de requisito tendrá un nombre, una descripción, y tres propiedades del requisito: importancia del requisito (alta, media o baja), prioridad del requisito (alta, media o baja) y complejidad del requisito (alta, media o baja).

TABLA 3.1. PLANTILLA DE TABLA DE REQUISITOS

RY-XX			
Nombre			
Descripción			
Necesidad	Alta	Media	Baja
Prioridad	Alta	Media	Baja
Complejidad	Alta	Media	Baja

3.1.1. Requisitos funcionales

TABLA 3.2. REQUISITO FUNCIONAL RF-01

RF-01	
Nombre	Captura de imágenes
Descripción	Capturar imágenes de los ocupantes en tiempo real.
Necesidad	Alta
Prioridad	Alta
Complejidad	Media

TABLA 3.3. REQUISITO FUNCIONAL RF-02

RF-02	
Nombre	Frecuencia de captura
Descripción	Ajustar la frecuencia de captura para optimizar rendimiento y precisión.
Necesidad	Media
Prioridad	Media
Complejidad	Media

TABLA 3.4. REQUISITO FUNCIONAL RF-03

RF-03	
Nombre	Calidad de captura
Descripción	Utilizar cámaras de alta resolución para la claridad de las imágenes.
Necesidad	Media
Prioridad	Media
Complejidad	Alta

TABLA 3.5. REQUISITO FUNCIONAL RF-04

RF-04	
Nombre	Pre-procesamiento de imágenes
Descripción	Realizar preprocesamiento para enfocar la imagen sobre la cara.
Necesidad	Alta
Prioridad	Media
Complejidad	Media

TABLA 3.6. REQUISITO FUNCIONAL RF-05

RF-05	
Nombre	Envío de imágenes
Descripción	Enviar imágenes en tiempo real al modelo de predicción.
Necesidad	Alta
Prioridad	Alta
Complejidad	Alta

TABLA 3.7. REQUISITO FUNCIONAL RF-06

RF-06	
Nombre	Ánalisis de imágenes
Descripción	Analizar las imágenes con el modelo de CNN pre-entrenado.
Necesidad	Alta
Prioridad	Alta
Complejidad	Alta

TABLA 3.8. REQUISITO FUNCIONAL RF-07

RF-07	
Nombre	Tipos de emociones
Descripción	Detectar emociones correspondientes
Necesidad	Alta
Prioridad	Alta
Complejidad	Alta

TABLA 3.9. REQUISITO FUNCIONAL RF-08

RF-08	
Nombre	API de resultados
Descripción	Comunicar el resultado de la predicción a la API de interacción.
Necesidad	Alta
Prioridad	Alta
Complejidad	Media

TABLA 3.10. REQUISITO FUNCIONAL RF-09

RF-09	
Nombre	Acciones según emoción
Descripción	Ejecutar acciones en respuesta al resultado obtenido.
Necesidad	Alta
Prioridad	Alta
Complejidad	Media

TABLA 3.11. REQUISITO FUNCIONAL RF-10

RF-10	
Nombre	Registro de acciones
Descripción	Registrar acciones realizadas y emociones detectadas.
Necesidad	Media
Prioridad	Media
Complejidad	Media

TABLA 3.12. REQUISITO FUNCIONAL RF-11

RF-11	
Nombre	Personalización de acciones
Descripción	Permite personalización de acciones por parte del usuario.
Necesidad	Baja
Prioridad	Media
Complejidad	Media

TABLA 3.13. REQUISITO FUNCIONAL RF-12

RF-12	
Nombre	Compatibilidad de vehículos
Descripción	El sistema tendrá que ser compatible con diferentes vehículos.
Necesidad	Alta
Prioridad	Media
Complejidad	Alta

TABLA 3.14. REQUISITO FUNCIONAL RF-13

RF-13	
Nombre	Integración con otros sistemas
Descripción	Integrar con otros sistemas de información.
Necesidad	Baja
Prioridad	Baja
Complejidad	Alta

TABLA 3.15. REQUISITO FUNCIONAL RF-14

RF-14	
Nombre	Mantenimiento remoto
Descripción	Soportar mantenimiento remoto y solución de problemas.
Necesidad	Media
Prioridad	Media
Complejidad	Media

3.1.2. Requisitos no funcionales

TABLA 3.16. REQUISITO NO FUNCIONAL RNF-01

RNF-01	
Nombre	Tiempo de respuesta
Descripción	El sistema deberá procesar la respuesta en tiempo real.
Necesidad	Alta
Prioridad	Alta
Complejidad	Alta

TABLA 3.17. REQUISITO NO FUNCIONAL RNF-02

RNF-02	
Nombre	Escalabilidad
Descripción	El sistema debe poder escalar para manejar múltiples usuarios a la vez.
Necesidad	Alta
Prioridad	Alta
Complejidad	Alta

TABLA 3.18. REQUISITO NO FUNCIONAL RNF-03

RNF-03	
Nombre	Robustez
Descripción	El sistema deberá poder manejar errores o fallos sin interrumpir el funcionamiento.
Necesidad	Alta
Prioridad	Alta
Complejidad	Alta

TABLA 3.19. REQUISITO NO FUNCIONAL RNF-04

RNF-04	
Nombre	Seguridad de datos
Descripción	Garantizar la seguridad de los datos de imágenes y emociones.
Necesidad	Alta
Prioridad	Alta
Complejidad	Alta

TABLA 3.20. REQUISITO NO FUNCIONAL RNF-05

RNF-05	
Nombre	Cumplimiento de normativa
Descripción	El sistema deberá cumplir la normativa de protección de datos y privacidad.
Necesidad	Alta
Prioridad	Alta
Complejidad	Media

TABLA 3.21. REQUISITO NO FUNCIONAL RNF-06

RNF-06	
Nombre	Eficiencia energética
Descripción	El sistema deberá ser eficiente en el consumo de energía.
Necesidad	Media
Prioridad	Media
Complejidad	Media

TABLA 3.22. REQUISITO NO FUNCIONAL RNF-07

RNF-07	
Nombre	Latencia
Descripción	El sistema debe tener una latencia de procesamiento mínima para asegurar una buena experiencia de usuario.
Necesidad	Alta
Prioridad	Alta
Complejidad	Alta

3.2. Análisis de datos

Como se ha mencionado anteriormente, el sistema obtiene los datos a partir de imágenes en tiempo real. Estas imágenes podrían ser recopiladas por cámaras instaladas dentro del vehículo, y posteriormente preprocesadas por Visión Artificial (Computer Vision) u otra técnica para obtener una imagen más detallada de la cara, y adaptadas al formato que se necesite, y será procesada posteriormente por el modelo de aprendizaje automático.

En este caso, para el entrenamiento del modelo se usará un conjunto de datos publicado en Internet, concretamente “FER+ (FER2013Plus)” [20] publicada por Microsoft en su Github [21], una versión extendida y mejorada del dataset original “Facial Expression Recognition (FER2013)” [22], publicado en Kaggle como enunciado de la competición “Challenges in Representation Learning: Facial Expression Recognition Challenge”.

El dataset FER2013 consta de 35.887 imágenes en escala de gris de caras humanas, con una resolución de 48 x 48 píxeles, adecuada con un balance entre calidad de visión y eficiencia computacional. Las imágenes tienen 7 categorías de emoción representadas en columnas y con valores numéricos dependiendo de la cantidad de presencia de esa emoción en la imagen, sumando un total de 10. La clasificación de las emociones son: enfado (anger), disgusto (disgust), miedo (fear), felicidad (happy), tristeza (sad), sorpresa (surprise) u neutral. Todos los datos de este conjunto se obtuvieron de la recopilación de imágenes de Internet y anotación manual por parte de personas que clasificaron las imágenes sobre las 7 categorías mencionadas anteriormente.

En cuanto a la versión extendida FER+, tiene mejoras del dataset original en relación a la reanotación de las imágenes de una manera más precisa, eliminando imágenes de baja calidad, cambiando anotaciones de imágenes incorrectamente anotadas, reduciendo el ruido y mejorando la calidad general del dataset. Además añade 3 categorías más para el grado de clasificación de las imágenes, que son desprecio (contempt), desconocido (unknown) y NT (not a face). Esta es la causa de la elección de este dataset, al contener una cantidad amplia de datos, con un etiquetado más preciso y mejorado que el original.



Fig. 3.2. Ejemplos de reanotación de imágenes (arriba FER2013 original, abajo FER+). Fuente: [21]

Todo lo comentado anteriormente es la causa de la elección de este dataset, al contener una cantidad amplia de datos, con un etiquetado más preciso y mejorado que el original. Además, en comparación con otros datasets como CK+ (Extended Cohn-Kanade) [23], que contiene 5.876 imágenes de expresiones faciales etiquetadas de 123 sujetos a color y en escala de grises, o el dataset JAFFE (Japanese Female Facial Expression) [24], que incluye 213 imágenes etiquetadas de expresiones faciales de 10 mujeres japonesas en escala de grises, FER2013Plus ofrece una mayor cantidad de datos y diversidad en las emociones representadas, además de estar todas las imágenes en escala de grises y no en dos tipos diferentes de formato, lo hace ser más preferible para entrenar modelos robustos y precisos en comparación con los otros dos, que aunque puedan ser también útiles, pensamos que no alcanzan el mismo nivel de cantidad de datos y estado en comparación para nuestro entrenamiento efectivo del modelo.

TABLA 3.23. TABLA DE COMPARACIÓN DE DATASETS

Dataset	Cantidad	Resolución	Estado
Fer2013Plus	35.887	48x48	Escala de grises
CK+	5.876	640x480	Color/Escala de grises
JAFFE	213	256x356	Escala de grises



Fig. 3.3. Ejemplos de imágenes del dataset CK+.



Fig. 3.4. Ejemplos de imágenes del dataset JAFFE.

3.3. Análisis de sistemas y software utilizado

Para la implementación y análisis del modelo de aprendizaje automático, se harán uso de diferentes programas informáticos y herramientas:

- **Python:** es un lenguaje de programación de alto nivel que destaca en el ámbito de la inteligencia artificial (IA) y el aprendizaje automático (ML) por su integración con una gran cantidad de librerías especializadas para el manejo de grandes volúmenes de datos, además de simplificar la creación, el entrenamiento y la evaluación de modelos predictivos de ML. [25]
- **Keras:** es una librería de Python que facilita el desarrollo de modelos de aprendizaje profundo con redes neuronales, además de proporcionar módulos para procesar imágenes, texto y secuencias de datos. [26]
- **NumPy y Pandas:** son dos librerías fundamentales que se utilizan para la interacción con datos numéricos. NumPy (Numerical Python) es destacado por una manipulación rápida y eficiente de grandes volúmenes de datos, soportando incluso datos en forma de arrays multidimensionales y matrices, y operación con una amplia cantidad de funciones matemáticas. [27] Pandas es una librería construida sobre NumPy, y destaca por una manipulación de datos eficiente con Series y DataFrames, estructuras unidimensionales la primera y bidimensionales la segunda, similar a columnas o tablas de una hoja de cálculo. También incluye funciones capaces de importar y exportar datos a diferentes tipos de archivos como Excel o CSV. [28]
- **Matplotlib y Seaborn:** son dos librerías centradas en la visualización de datos. Matplotlib destaca por un control detallado sobre la creación de distintos tipos de gráficos simples y complejos, proporcionando flexibilidad para modificar prácticamente todos los aspectos de los gráficos, pero a costa de mayor complejidad. [29] Seaborn por otro lado es una librería construida sobre Matplotlib, que simplifica la creación de gráficos y visualizaciones, y facilita la exploración y visualización de los datos. [30]

4. IMPLEMENTACIÓN Y RESULTADOS

En este apartado se detalla el proceso de creación y evaluación del modelo de reconocimiento de emociones faciales, divide en tres secciones principales. La primera sección, "Preprocesado de datos", describe el conjunto de datos utilizado (FER+), su distribución inicial y los pasos tomados para balancear las clases de emociones. La segunda sección, "Entrenamiento del modelo", explica el proceso de entrenamiento de la red neuronal convolucional (CNN), incluyendo la configuración de los parámetros del modelo. Finalmente, la tercera sección, "Análisis de resultados", presenta una evaluación del rendimiento del modelo, utilizando las gráficas de rendimiento y la matriz de confusión.

4.1. Preprocesado de datos

Como se ha comentado anteriormente, el dataset que se usará es el FER+. El dataset tiene las imágenes etiquetadas y clasificadas en conjuntos de entrenamiento, validación y test. Tras descargar los datos y realizar un análisis, se obtiene la siguiente distribución de valores:

TABLA 4.1. TABLA DE DISTRIBUCIÓN INICIAL DE VALORES

Emoción	Cantidad
Neutral	12.906
Happiness	9.355
Surprise	4.462
Sadness	4.371
Anger	3.111
Disgust	248
Fear	819
Contempt	216
Unknown	222
Not a face	4
Total	35.714

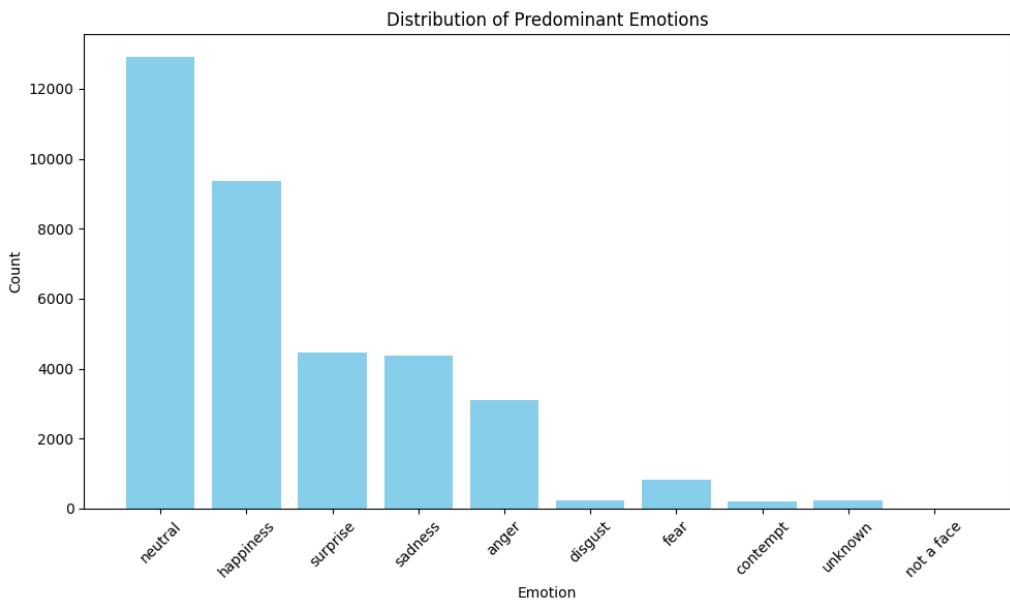


Fig. 4.1. Gráfico de distribución inicial de valores.

Se puede apreciar que los datos son muy desbalanceados, los datos que más hay son de etiqueta “neutral” y “happiness”, mientras que de “not a face” hay solo 4. Se decide que los datos de “not a face” y “unknown” se desechan ya que no son relevantes y no aportan valor. Para balancear los datos, también se decide desechar “disgust”, “contempt” y “fear”, ya que en comparación también son datos bastante desbalanceados, además de que son parecidos a “sadness”, “anger” y “surprise”. Por último, se eliminan entradas de “neutral” y “happiness” ya que tienen un número de entradas mucho mayor que los demás, quedando al final con una cantidad de entre 3.000 y 5.000:

TABLA 4.2. TABLA DE DISTRIBUCIÓN BALANCEADA DE VALORES

Emoción	Cantidad
Neutral	4.906
Happiness	4.355
Surprise	4.462
Sadness	4.371
Anger	3.111
Total	21.205

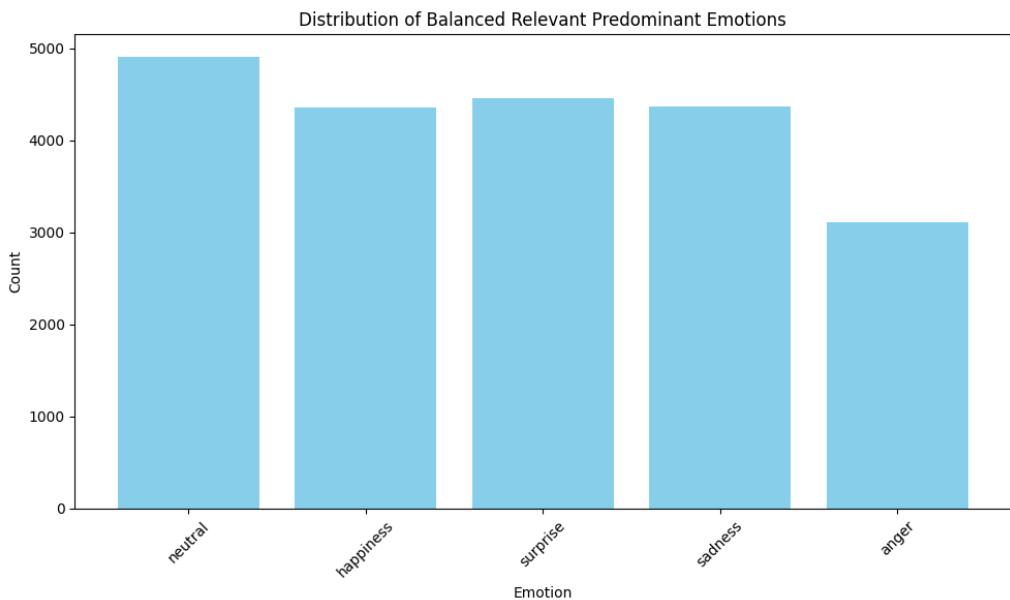


Fig. 4.2. Gráfico de distribución balanceada de valores.

También se muestra el número de valores de cada tipo de cada conjunto clasificado en entrenamiento, validación y test:

TABLA 4.3. TABLA DE COMPARACIÓN DE DIFERENTES DATASETS

Conjunto	Neutral	Happines	Surprise	Sadness	Anger	Total
Entrenamiento	4.009	3.428	3.562	3.515	2.467	16.981
Validación	435	499	456	412	319	2.121
Test	462	428	444	444	325	2.103
Total						21.205

Por último, antes de realizar el entrenamiento del modelo con la CNN se normalizan los píxeles de las imágenes a valores entre 0 y 1 para ayudar al modelo a converger mejor.

4.2. Entrenamiento del modelo

El entrenamiento del modelo se ha realizado en un cuaderno Juoyter, disponible en la siguiente página de Github: <https://github.com/QianXunX/TFG>

Tras el preprocesado de los datos, se procede al entrenamiento de un modelo de red neuronal convolucional (CNN) con los datos. La red neuronal creada será una red convolucional con 10 capas en total, que se diferencian en:

- **Capa Conv2D:** son capas convolucionales que aplican un determinado número de filtros a la imagen, y cada filtro extrae diferentes características de la imagen de en-

trada, como bordes, colores o formas, cada vez más complejas según la profundidad de la red. Las capas también tienen un tamaño de kernel, el tamaño de la ventana que desliza sobre las imágenes para crear el mapa de características de la imagen. En este caso se hace uso de filtros de 32 para detectar características básicas como bordes y texturas simples, filtros de 64 captaron características más complejas, y filtros de 128 para extraer características de alto nivel como formas y patrones complejos. El kernel usado es de 3x3, ya que es un tamaño para capturar suficientes relaciones sin que sea demasiado costoso computacionalmente. La función de activación usada en estas capas convolucionales es la Unidades Lineales Rectificadas (ReLU) para introducir no linealidad en el modelo y poder aprender patrones más complejos.

- **Capa MaxPooling2D:** son capas que reducen la dimensionalidad de la imagen (ancho y alto) para la siguiente capa convolucional, para ayudar al modelo a detectar características sin ser afectadas por pequeños cambios de escala y orientación, además de disminuir el coste computacional y el sobreajuste. En este caso se usa un tamaño de 2x2, un equilibrio entre la reducción de dimensionalidad sin que se pierda mucha información.
- **Capa Flatten:** son capas que aplanan la salida de las capas convolucionales a vectores de una sola dimensión, necesario para la entrada de datos de las capas densas. Esta capa es esencial para proporcionar la información a las capas densas de manera eficiente.
- **Capa densa:** son capas completamente conectadas que clasifican la imagen basado en las características extraídas por las capas convolucionales. También se usa Dropout, una técnica que sirve para eliminar unidades de salida de la capa de manera aleatoria para prevenir sobreajuste del modelo. En este caso se usa un Dropout de 0.5, eliminando la mitad de las salidas, lo que fuerza a la red a aprender representaciones redundantes y robustas de las características.
- **Capa de salida:** la última capa de la red, en este caso una capa densa utilizando una función de activación Softmax, que sirve para sumar 1 a la salida y poder interpretarse como valores de probabilidad. En este caso esta capa de salida tiene 5 unidades, correspondientes a cada una de las 5 emociones que el modelo clasifica.

Tras definir el modelo de CNN, es necesario compilarlo para configurar diferentes opciones de entrenamiento, como son el optimizador, la función de pérdida y las métricas

1. **Optimizador:** son algoritmos o métodos responsables de cambiar los atributos de las redes neuronales, como los pesos o la tasa de aprendizaje, con el fin de minimizar la función de pérdida. Por lo tanto, ayudan a minimizar o maximizar una función objetivo (función de pérdida) actualizando de manera iterativa los parámetros del modelo.

- Descenso de Gradiente Estocástico (SGD): es simple y básico que actualiza los pesos de manera iterativa, pero depende mucho de la tasa de aprendizaje.
- Algoritmo de Gradiente Adaptativo (AdaGrad): método que mantiene la tasa de aprendizaje en base a la frecuencia de actualización de cada parámetro mejorando el rendimiento en problemas con gradientes dispersos. La tasa de aprendizaje no requiere ajuste manual, pero puede volverse muy pequeña y detener el aprendizaje.
- Propagación de Promedio Cuadrático (RMSprop): método que también mantiene tasas de aprendizaje adaptativas para cada parámetro basado en el promedio de las magnitudes más recientes. Establece el entrenamiento con tasas de aprendizaje más consistentes, pero requiere un buen ajuste de hiperparámetros.
- Estimación de Momento Adaptativo (Adam): método que combina las ventajas de RMSProp y AdaGrad por lo que mantiene tasas de aprendizaje adaptativos para cada parámetro, pero puede ser computacionalmente más costoso que SGD. Aun así este es el optimizador que se usará al aportar más ventajas que las demás.

2. **Función de pérdida:** es la función objetivo que el modelo intenta minimizar durante el entrenamiento, y significa qué tan bien el modelo preside la variable objetivo. Depende del tipo de datos que se entrena y la tarea, si es clasificación, regresión, etc. Funciones de pérdida comunes son “categorical crossentropy” para tareas de clasificación multiclas mutuamente exclusivas, “binary crossentropy” para tareas de clasificación binaria, “mean square error (MSE)” para tareas de regresión, etc. En este caso se usará “categorical crossentropy” al ser un problema de clasificación con múltiples clases mutuamente excluyentes.

3. **Métricas:** valores para evaluar el rendimiento del modelo, parecido a las funciones de pérdida, pero los resultados de la evaluación no se utilizan al entrenar el modelo. Un ejemplo es la precisión (accuracy), el valor que mide la frecuencia que las predicciones son iguales a los valores etiquetados, utilizada en tareas de clasificación, como en este caso.

Para la opciones de compilación, el optimizador elegido es Adam, ya que como se ha comentado anteriormente combina las ventajas de dos optimizadores (AdaGrad y RMSProp), siendo más eficaz que los otros y requerir menos memoria de cómputo. La función de pérdida que se usará es una función “categorical crossentropy” que mide la diferencia entre la distribución de probabilidad de los valores predichos con la de los valores reales, adecuado para problemas de clasificación multiclas como en este caso. La métrica usada es la de precisión, al ser la más intuitiva y clara para medir el porcentaje de predicciones correctas respecto al total. Todas estas opciones se han elegido para optimizar el rendimiento y la eficacia del modelo de CNN.

4.3. Análisis de resultados

El entrenamiento de este modelo resulta una precisión de un 70.14 % de emociones predichas correctamente. En comparación con otros estudios de investigación que han utilizado el mismo dataset, como por ejemplo el estudio “Auto-FERNet: A Facial Expression Recognition Network With Architecture Search” [31] donde diseñaron un modelo de red que obtuvo una precisión de un 73.78 %, o por ejemplo en el estudio “Hybrid Facial Expression Recognition (FER2013) Model for Real-Time Emotion Classification and Prediction” [32] donde también diseñaron una red CNN con la que obtuvieron una precisión de 70 %, se puede apreciar que a pesar de la variación en la arquitectura y las técnicas del modelo, los resultados de precisión son similares, lo que indica que el resultado obtenido está dentro del rango esperado para este dataset, pero que todavía es mejorable ampliando el diseño con el uso de diferentes otras técnicas.

Las siguientes gráficas muestran los valores de la función de pérdida y precisión a lo largo de las rondas de entrenamiento.

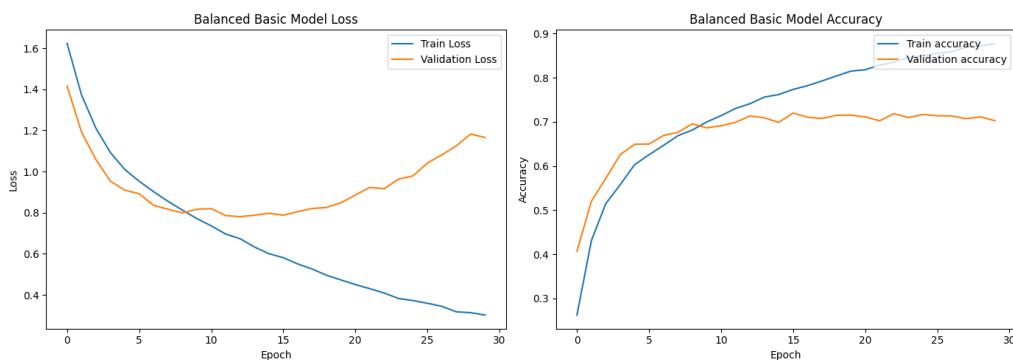


Fig. 4.3. Gráficas de valores de función de pérdida y valores de precisión.

Si observamos la primera gráfica (izquierda), la gráfica de valores de la función de pérdida, se puede apreciar que la pérdida del entrenamiento (azul) disminuye de manera constante, lo que significa que el modelo está aprendiendo correctamente y mejorando en cada iteración. En cambio, la pérdida de la validación (naranja) disminuye en un inicio pero comienza a estabilizarse y aumentar ligeramente después de aproximadamente 15 iteraciones. Esto podría indicar un sobreajuste del modelo, significando que el modelo aprende ruido y detalles no aplicables a los datos generales.

Observando la segunda gráfica (derecha), la gráfica de valores de precisión, se puede apreciar que la precisión del entrenamiento mejora de manera consistente en el tiempo, lo cual es un resultado esperado ya que el modelo se ajusta cada vez más a los datos de entrenamiento. Pero la precisión con los datos de validación, al igual que la función de pérdida, aumenta en un inicio pero posteriormente se estabiliza, lo que también puede indicar indicios de sobreajuste.

La siguiente imagen de la matriz de confusión muestra la cantidad de valores clasifi-

cados correctamente por el modelo cada emoción en comparación con los valores etiquetados reales.

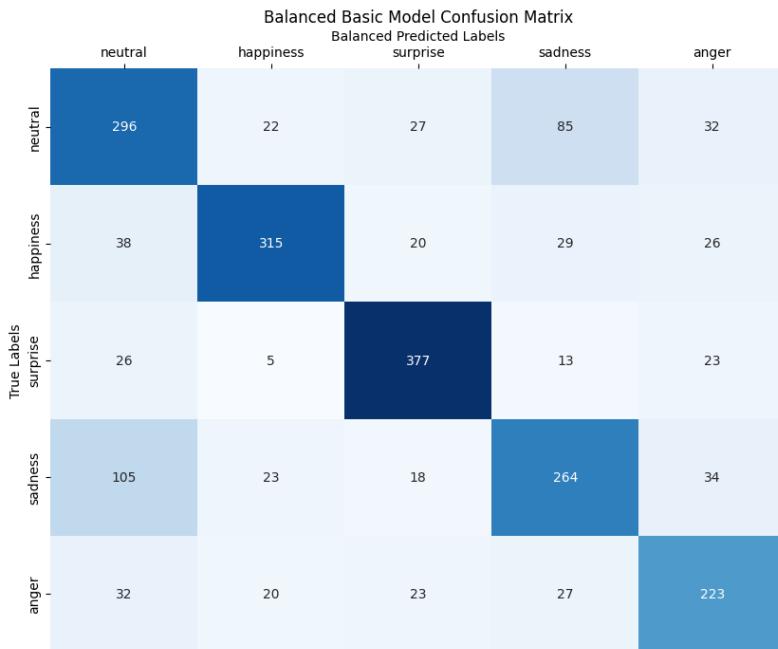


Fig. 4.4. Matriz de confusión de predicciones del modelo.

Las entradas que aparecen en la diagonal, verdaderos positivos, representan el número de predicciones correctas para cada tipo de emoción. Se puede apreciar que la mayoría de los valores se predicen de manera correcta, de esto el valor de acierto de un 70.14

- Neutral: 296 predicciones correctas.
- Happiness: 315 predicciones correctas.
- Surprise: 377 predicciones correctas.
- Sadness: 264 predicciones correctas.
- Anger: 223 predicciones correctas.

En cambio, las entradas que no están en la diagonal, falsas clasificaciones, representan predicciones incorrectas del modelo en función de cada valor. Las predicciones incorrectas entre las clases “neutral” y “sadness” muestran un número significativo de valores, siendo clasificadas 105 de “sadness” como “neutral” y 85 de “neutral” como “sadness”. Esto puede ser debido a que ambas emociones comparten características visuales similares, y siendo la resolución de datos baja (48x48 píxeles en escala de grises), los detalles que diferencian ambas emociones sean difíciles de capturar y distinguir para el modelo, especialmente en aquellos casos donde la expresión facial de la emoción “sadness” sea mínima. Aunque el número de clasificaciones incorrectas no sean tantas (un tercio) en

comparación con las correctas, una mejora en la resolución o con imágenes en donde la emoción de “sadness” sea más diferenciada se podría obtener una mayor precisión.

Cabe comentar que en la predicción de las emociones “happiness” y “anger” el modelo se desempeña de manera positiva, dado que ambas emociones son contradictorias, una interpretación precisa entre ellas es crucial ya que respuestas o acciones basadas en emociones contradictorias pueden llevar a experiencias negativas. Un número reducido de predicciones erróneas entre ambas (aproximadamente 10 %) en comparación con las correctas indica que el modelo es capaz de identificar de manera efectiva entre ambos estados emocionales contrastantes.

En general, el modelo con el dataset FER+ ha obtenido un resultado de predicción no excelentemente alto pero razonablemente bueno, siendo un 70.14 % teniendo en cuenta que las imágenes de este dataset eran de una baja calidad y resolución, de 48x48 píxeles y en escala de grises, lo que hace el problema más difícil en comparación con datos de imágenes de alta resolución.

Se considera que lo más importante sería considerar una mejora o incremento en la calidad de imagen, o uso de técnicas de preprocesamiento que resalten características importantes de las emociones de una manera más clara para que el modelo pueda aprender mejor. También se podría considerar usar redes pre-entrenadas con grandes cantidades de datos similares a las de el dataset utilizado, ya que el modelo pre-entrenado tendría más parámetros y características aprendidas para predecir mejor. Otra opción sería realizar más preprocesamiento sobre los datos, como por ejemplo un aumento de los datos mediante técnicas de rotación o translación, para tener una variedad de datos mayor y mejorar la capacidad de generalización del modelo.

5. MARCO REGULADOR E IMPACTO SOCIO-ECONÓMICO

5.1. Marco regulador

En cuanto al marco regulador de este proyecto, en el proceso de desarrollo del sistema de reconocimiento de emociones en los vehículos, hay que tener en cuenta varias normativas y leyes para que se garantice la legalidad y seguridad de dichos datos personales. Una de las normativas más importantes y relevantes para este proyecto es la Ley Orgánica 3/2018 de Protección de Datos Personales y Garantía de los Derechos Digitales (LOPDGDD) [33].

La Ley Orgánica 3/2018 es una adaptación y complementación al Reglamento General de Protección de Datos (RGPD) de la Unión Europea [34], en la que se establecen normas sobre la recopilación, el almacenamiento y el uso de los datos personales. En relación a este proyecto, implica que cualquier dato obtenido a través del sistema de reconocimiento de emociones debe tener el nivel de seguridad y privacidad más alto. Esto es debido a que los datos faciales de las personas son datos muy sensibles, y se necesita el consentimiento de los usuarios para su captación, procesamiento y uso. Además, también se tienen que implementar medidas de seguridad para proteger dichos datos de accesos no autorizados y garantizar que no sean usados para otros fines distintos a los de este proyecto.

Otra normativa relacionada es el Real Decreto 662/2012 [35], que establece normas para la implantación de los sistemas inteligentes de transporte (SIT) y promueve la integración de estas tecnologías para mejorar la seguridad y la eficiencia. En relación a nuestro proyecto, la normativa establece que nuestro diseño del sistema de reconocimiento de emociones debe alinearse con los principios de SIT, y asegurar que su implementación no solo mejora la experiencia de usuario, sino que también contribuye a la seguridad en las vías, incluyendo que el sistema debe ser funcional y operable con otros SIT, y cumplir con los estándares definidos en el Real Decreto.

También existe una Ley de IA de la UE [36], todavía en desarrollo, siendo la primera normativa específica sobre el uso de tecnologías de Inteligencia Artificial en Europa. Esta ley introduce requisitos de transparencia, aplicabilidad y responsabilidad para los sistemas IA, que influye en nuestro proyecto como la transparencia en el funcionamiento del sistema, permitiendo a los usuarios entender la manera a la que se procesan sus datos y cómo se toman las decisiones basadas en el análisis de emociones. Además, el sistema debería incluir mecanismos de responsabilidad permitiendo identificar y corregir cualquier error de su funcionamiento, asegurando que el uso del mismo sea de manera ética y justa.

Tras este análisis del marco regulatorio, se sacan conclusiones acerca del cumplimiento normativo de nuestro proyecto. Se deberá de desarrollar políticas claras de privacidad e información sobre el uso de datos para los usuarios del sistema, además de medidas de

seguridad ya sea cifrado de datos o restricción de acceso a los mismos. También se deberá realizar auditorías para asegurar el cumplimiento de las normativas actuales, y adaptarse a cualquier posible cambio.

5.2. Impacto socio-económico

La implementación de este sistema IA en el ámbito económico puede promover la competitividad entre las diferentes empresas, ya que con este sistema podrían ofrecer una experiencia de usuario mejor, y por lo tanto tener posibilidad de aventajarse y atraer consumidores que busquen una mejor experiencia de usuario. Además, al igual que la creación de cualquier sistema, aumentaría la demanda de elementos hardware y software, así como los servicios de mantenimiento de estos sistemas.

Pero no todo es positivo, el uso de este sistema también podría tener efectos económicos negativos, ya sean los costes de desarrollo e implementación iniciales como el de mantenimiento posterior, podría ser difícil de lograr para empresas pequeñas o startup.

En relación al impacto social, es también significativo, ya que mejora la seguridad y el confort de los pasajeros. Desde el punto de vista de la experiencia de usuario, puede reducir el estrés de los pasajeros y aumentar su satisfacción durante un viaje. Desde el punto de vista de seguridad, es posible usarse para detectar emociones que puedan afectar a la seguridad de los pasajeros, y por lo tanto a otros usuarios de las vías, o ayudar a actuar de manera autónoma al vehículo.

Pero por otro lado, también tiene que asegurarse que tanto la privacidad como el uso ético de los datos captados sea gestionado de manera segura y solo para los fines que se especifican en este proyecto. Esta transparencia en el uso de los datos y en el funcionamiento del sistema es la parte esencial para obtener confianza y que el público lo acepte.

En relación al impacto medioambiental, puede ser positivo si se utiliza para optimizar el rendimiento y reducir el consumo de combustible, por ejemplo ajustando el sistema de calefacción de manera más eficiente. Pero puede ser todo lo contrario, que aumente el gasto al usar más calefacción de lo normal, siendo en este caso implementado en vehículos no eléctricos. Si se implementara en vehículos eléctricos no habría este problema, y también sería un beneficio medioambiental si esto fomenta a más personas a comprar vehículos eléctricos.

Y en cuanto al impacto ético, como se ha mencionado anteriormente, la principal preocupación es la privacidad y el uso de los datos generados. En este caso es fundamental que se implementen políticas estrictas de protección de datos y que esos datos sólo sean usados para usos transparentes mencionados en este proyecto.

6. GESTIÓN DEL PROYECTO

En este apartado se abordará la planificación y estimación de los recursos necesarios para llevar a cabo el proyecto. Este apartado incluye un diagrama de Gantt que detalla la planificación de las actividades a lo largo de un período de aproximadamente 3.5 meses. Además, se describen los roles del personal requerido, junto con una tabla de costes de personal basada en los salarios anuales de cada posición. También se incluyen los costos de hardware y software necesarios para el proyecto, calculados en función de su amortización durante la duración del mismo.

6.1. Planificación

El siguiente diagrama de Gantt detalla la planificación de este proyecto, resultando un total de 72 días, aproximadamente 3.5 meses. En este proyecto se han abordado las siguientes tareas clave:

1. Definición de proyecto y objetivos
2. Investigación y análisis inicial
3. Diseño del sistema e identificación de requisitos
4. Planificación y estimación del proyecto
5. Preprocesado de datos
6. Entrenamiento de un modelo
7. Análisis de resultados
8. Evaluación final de cumplimiento normativo e impacto socioeconómico

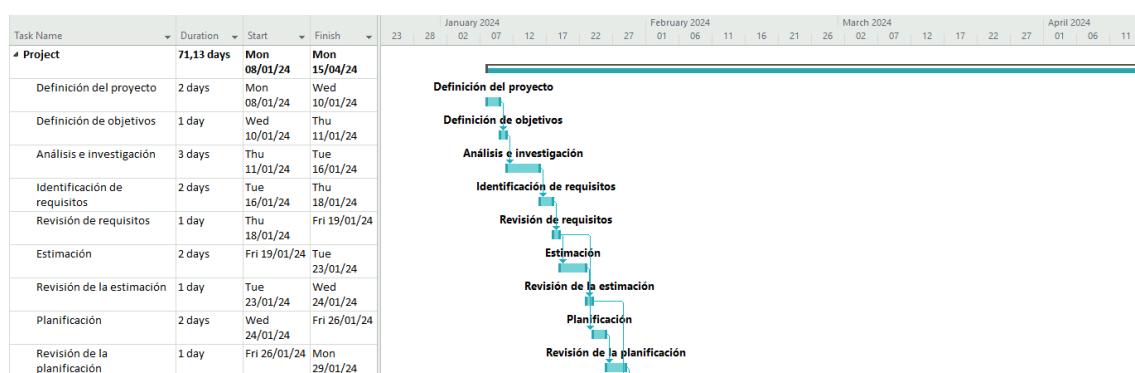


Fig. 6.1. Diagrama de Gantt. (Parte 1)

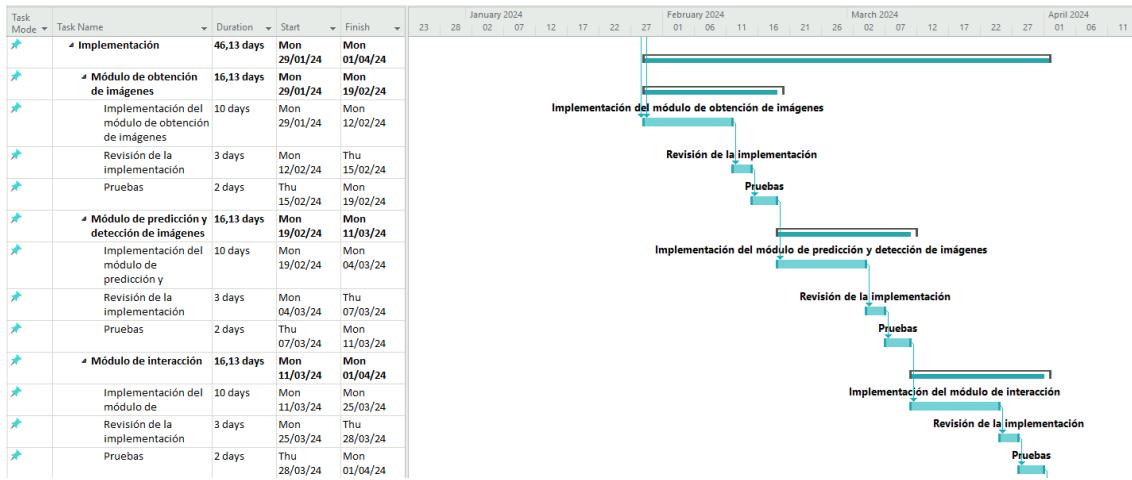


Fig. 6.2. Diagrama de Gantt. (Parte 2)



Fig. 6.3. Diagrama de Gantt. (Parte 3)

6.2. Estimación

6.2.1. Costes personal

Para este proyecto se necesitarán los siguientes roles de personal:

- **Arquitecto de software:** responsable de tomar decisiones de alto nivel en la arquitectura del software, y guiar su diseño y desarrollo.
- **Ingeniero de ML:** responsable de desarrollar y entrenar el modelo de red neuronal para la detección de emociones, y optimizar su desempeño.
- **Científico de datos:** responsable de analizar los datos y mejorar los algoritmos utilizados, trabajar en el preprocesamiento de los datos antes del entrenamiento.
- **Ingeniero de ciberseguridad:** responsable de desarrollar estrategias para proteger los sistemas, redes y datos contra posibles amenazas o ataques.
- **Ingeniero de sistemas:** responsable de diseñar, implementar y mantener la infraestructura software.
- **QA Tester:** responsable de realizar pruebas para asegurar la calidad y funcionalidad del sistema, identificando y reportando posibles errores.

- **Project Manager:** responsable de planificación, ejecución y control del proyecto, asegurando que se cumplan los objetivos del presupuesto y tiempo establecidos.

TABLA 6.1. TABLA DE COSTES DE PERSONAL

Posición	Cantidad	Salario (€/año)	Subtotal (€/3.5 meses)
Arquitecto de software	1	38.000	11.083
Ingeniero de ML	2	40.000	11.666 * 2 = 23.332
Científico de datos	2	40.000	11.666 * 2 = 23.332
Ingeniero de ciberseguridad	2	35.000	10.208 * 2 = 20.416
Ingeniero de sistemas	2	28.000	8.166 * 2 = 16.332
QA Tester	2	27.000	7.875 * 2 = 15.750
Project Manager	1	48.000	14.000
Coste total personal			124.245

Los datos de salarios se han obtenido de la guía salarial Talentfy de perfiles IT de España del primer semestre de 2024. [37]

6.2.2. Costes hardware

El coste final de los recursos hardware se calcula en base al coste total inicial del producto dividiendo por el tiempo de vida de ese producto y posteriormente multiplicando por la duración del proyecto, resultando el gasto de amortización del recurso hardware, lo que es el coste final.

TABLA 6.2. TABLA DE COSTES HARDWARE

Recurso	Cantidad	Coste inicial (€/unidad)	Coste final (€)
Portátiles de empresa	12	1.329	$96,9 * 12 = 1.162,8$
Cámaras	4	59,49	$8,68 * 4 = 34,72$
Servidor de entrenamiento	1	3.577	260,82
Almacenamiento	4	19,99	$2,92 * 4 = 11,68$
Otros accesorios y perifericos	1	1.000	145,83
Coste total hardware			1.615,85

El tiempo de vida de los PCs son de aproximadamente 4 años, y de las cámaras y los accesorios 2 años. Por lo que el coste final se calcula multiplicando por su tiempo de amortización, 3.5 meses/48 meses y 3.5 meses/ 24 meses respectivamente.

6.2.3. Costes software

El coste software dependería solo de las posibles licencias de software que se puedan a llegar a usar en el desarrollo del proyecto. En este caso se podría considerar la licencia para el IDE Pycharm de Jetbrains con un coste anual de 249€ por cada empleado, y dando un coste inicial de $249 * 12 = 2.988\text{€}$ y un coste de amortización de $2.988 / 12 * 3,5 = 871,5\text{€}$.

También se podría considerar la licencia de StarUML para el arquitecto de software, con un coste inicial anual de 122 € y un coste de amortización de $122 / 12 * 3,5 = 35,58\text{€}$.

Por lo tanto el coste total de software final sería de 907,08€.

6.2.4. Resumen de costes

En la siguiente tabla se muestra el coste total de cada parte, además del coste total final. Tener en cuenta que en la tabla también se muestran costes aproximados de materiales fungibles que puedan ser utilizados como por ejemplo hojas de impresora, bolígrafos, carpetas, etc.

TABLA 6.3. TABLA DE RESUMEN DE COSTES

Descripción	Total
Salario de empleados	124.245€
Equipamiento hardware y software	1.615,85€
Fungibles	135€
Total (sin IVA)	125.995,85€
Beneficios (10 %)	12.599,56€
Margin de riesgos (15 %)	18.899,38€
Total (sin IVA) con beneficios y riesgos +15 %	157.494,79€
Total (sin IVA) con beneficios y riesgos -15 %	119.696,03€
VAT(21 %) con beneficios y riesgos +15 %	57.958,07€
VAT(21 %) con beneficios y riesgos -15 %	20.159,31€
Total (con IVA) con beneficios y riesgos +15 %	183.953,92€
Total (con IVA) con beneficios y riesgos -15 %	146.155,16€

7. CONCLUSIONES Y TRABAJOS FUTUROS

El desarrollo e implementación de un sistema de reconocimiento de emociones faciales en vehículos ha demostrado ser una herramienta valiosa tanto para mejorar la experiencia del usuario como para aumentar la seguridad en la conducción. A través del análisis y procesamiento de datos en tiempo real, el sistema podría ajustar dinámicamente diversos aspectos del ambiente del vehículo, como la música, la iluminación y la temperatura, ofreciendo una experiencia más personalizada y confortable para los pasajeros.

En términos generales, los objetivos iniciales planteados han sido cumplidos satisfactoriamente. Se ha desarrollado un modelo de reconocimiento de emociones faciales que ha alcanzado una precisión razonable del 70.14 % utilizando el dataset FER+. También se ha diseñado e implementado un sistema que integra el modelo con los sistemas del vehículo, permitiendo la interacción con controles como la iluminación y el audio. Además, se ha realizado un análisis de viabilidad económica y técnica, concluyendo que la integración de esta tecnología es factible y presenta un impacto positivo en la experiencia del usuario y en la seguridad de los pasajeros. No obstante, también es importante considerar los posibles costes iniciales de desarrollo e implementación, que pueden ser una barrera para empresas más pequeñas. Además, la privacidad y seguridad de los datos son aspectos críticos que deben ser gestionados con políticas estrictas y transparentes para asegurar la confianza del usuario.

La realización de este proyecto ha sido una experiencia enriquecedora tanto a nivel profesional como personal. La aplicación de conocimientos adquiridos a lo largo de la carrera en Ingeniería Informática ha sido fundamental para abordar los diversos desafíos técnicos y metodológicos que surgieron. Asignaturas clave como Inteligencia Artificial, Aprendizaje Automático me han aportado los conocimientos esenciales para la realización de este proyecto, y la asignatura de Dirección de Proyectos de Desarrollo Software ha sido fundamental para la planificación y organización de este proyecto.

El sistema desarrollado es solo una simple idea de diseño, que se podría mejorar de muchas maneras y realizar futuros trabajos en relación:

1. **Mejora de la Calidad de Datos:** implementar técnicas avanzadas de preprocesamiento de imágenes y utilizar datasets de mayor resolución para mejorar la precisión del modelo. Incluir clasificaciones con más emociones y características más claras para una mejor precisión en la detección.
2. **Integración con ADAS:** desarrollar funcionalidades adicionales que integren el reconocimiento de emociones con los sistemas avanzados de ayuda a la conducción para una experiencia de usuario más completa y segura, o incluir en sistemas ya existentes para diferentes propósitos.

3. **Evaluación en Entornos Reales:** realizar pruebas exhaustivas en escenarios de uso real para validar el desempeño del sistema en condiciones diversas y dinámicas.
4. **Optimización y Eficiencia:** investigar métodos para optimizar el modelo y reducir el coste computacional, facilitando su implementación en tiempo real y en dispositivos con recursos limitados.

En conclusión, el presente trabajo no solo abre nuevas posibilidades para la personalización y seguridad en los vehículos, sino que también sienta las bases para futuras innovaciones en la integración de inteligencia artificial y aprendizaje automático en la industria automotriz.

BIBLIOGRAFÍA

- [1] P. Papadaki, *User experience in automotive industry: user perspective on functionality and entertainment*, ene. de 2019. [En línea]. Disponible en: <https://www.diva-portal.org/smash/record.jsf?dswid=-6939&pid=diva2%3A1385278>.
- [2] M. T. Zulkefli, N. I. A. Shafai y H. A. Azizan, *Automotive Interior: A study on the Dashboard Touch Screen Panel and Its Impact to the Driver*. [En línea]. Disponible en: <https://eudl.eu/doi/10.4108/eai.24-8-2021.2315303>.
- [3] R. Coppola y M. Morisio, “Connected car: technologies, issues, future trends,” *ACM Computing Surveys (CSUR)*, vol. 49, n.º 3, pp. 1-36, 2016. [En línea]. Disponible en: https://dl.acm.org/doi/abs/10.1145/2971482?casa_token=Cj1Q3I4upV4AAAAA:HTAewkA3Unk064rhoeUvpV37pH-aexXDTPc3LeqMtX-HAALUXUAnQ6AloA
- [4] M. Maroto, E. Caño, P. González y D. Villegas, *Head-up Displays (HUD) in driving*, mar. de 2018. [En línea]. Disponible en: <https://arxiv.org/abs/1803.08383>.
- [5] M. i. DGT, *Sistemas avanzados de ayuda a la conducción (ADAS)*. [En línea]. Disponible en: <https://www.dgt.es/muevete-con-seguridad/conviertete-en-un-buen-conductor/Sistemas-avanzados-de-ayuda-a-la-conduccion-ADAS-/>.
- [6] *Los ADAS obligatorios desde julio 2022, uno a uno*. [En línea]. Disponible en: <https://revista.dgt.es/es/motor/tecnologia-seguridad/2021/0518-Landing-ADAS.shtml#>.
- [7] D. Parekh et al., “A Review on Autonomous Vehicles: Progress, Methods and Challenges,” *Electronics*, vol. 11, n.º 14, jul. de 2022. [En línea]. Disponible en: <https://www.mdpi.com/2079-9292/11/14/2162>.
- [8] *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. [En línea]. Disponible en: https://www.sae.org/standards/content/j3016_202104/.
- [9] F. Scurt, T. Vesselenyi, R. Tarca, H. Beles y G. Dragomir, “Autonomous vehicles: classification, technology and evolution,” en *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, vol. 1169, 2021, p. 012032. [En línea]. Disponible en: <https://iopscience.iop.org/article/10.1088/1757-899X/1169/1/012032/meta>.
- [10] *Advanced Driver-Assistance Systems: A Path Toward Autonomous Vehicles*. [En línea]. Disponible en: <https://ieeexplore.ieee.org/document/8429957>.

- [11] *Artificial intelligence, machine learning and deep learning.* [En línea]. Disponible en: https://ieeexplore.ieee.org/abstract/document/8259629?casa_token=e8Kq68ejBNMAAAA:ozsksy20L0JWpg0P9wKd6eJVqA9aJHVKc6ifX3a2rRMJ8yBZJ0-J0qGc2WNvkIW1jTVx5Q2icg.
- [12] R. I. Mukhamediev et al., “Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges,” *Mathematics*, vol. 10, n.º 15, jul. de 2022. [En línea]. Disponible en: <https://www.mdpi.com/2227-7390/10/15/2552>.
- [13] M. G. M. Abdolrasol et al., “Artificial Neural Networks Based Optimization Techniques: A Review,” *Electronics*, vol. 10, n.º 21, nov. de 2021. [En línea]. Disponible en: <https://www.mdpi.com/2079-9292/10/21/2689>.
- [14] PouyanfarSamira, “A Survey on Deep Learning: Algorithms, Techniques, and Applications: ACM Computing Surveys: Vol 51, No 5,” *ACM Computing Surveys (CSUR)*, sep. de 2018. [En línea]. Disponible en: https://dl.acm.org/doi/abs/10.1145/3234150?casa_token=NwDv6OKQHucAAAAA:Z-9c53n0-avqDPOCU_ujAByyBz1Zioe4FKmIOMsHhJjPViPxS_H6_h9B0YxV5ET6TkXAdf0HaL0g.
- [15] D. Zhou, Y. Cheng, L. Wen, H. Luo e Y. Liu, “Drivers’ Comprehensive Emotion Recognition Based on HAM,” *Sensors*, vol. 23, n.º 19, oct. de 2023. [En línea]. Disponible en: <https://www.mdpi.com/1424-8220/23/19/8293>.
- [16] H. Abedi, S. Luo, V. Mazumdar, M. M. Riad y G. Shaker, “AI-powered in-vehicle passenger monitoring using low-cost mm-wave radar,” *IEEE Access*, vol. 10, pp. 18 998-19 012, 2021. [En línea]. Disponible en: <https://ieeexplore.ieee.org/abstract/document/9666842>.
- [17] S. M. Asad, J. Ahmad, S. Hussain, A. Zoha, Q. H. Abbasi y M. A. Imran, “Mobility Prediction-Based Optimisation and Encryption of Passenger Traffic-Flows Using Machine Learning,” *Sensors*, vol. 20, n.º 9, mayo de 2020.
- [18] T. Laakkona, J. Kostiainen y O. Pihlajamaaa, “Estimating passenger experience from vehicle data: Preconditions for using machine learning,” *Proceedings of TRA2020, the 8th Transport Research Arena, Liikenne-ja viestintävirasto Traficom, Finland*, 2020. [En línea]. Disponible en: https://www.researchgate.net/profile/Timo-Laakko/publication/339848623_Estimating_passenger_experience_from_vehicle_data_Preconditions_for_using_machine_learning/_links/5e8efcc94585150839cae10e/Estimating-passenger-experience-from-vehicle-data-Preconditions-for-using-machine-learning.pdf.
- [19] M. Skhosana y A. E. Ezugwu, “A Real-Time Machine Learning-Based Public Transport Bus-Passenger Information System.,” *International Journal of Sustainable Development & Planning*, vol. 16, n.º 7, 2021. [En línea]. Disponible en: https://www.researchgate.net/profile/Absalom-Ezugwu/publication/357088862_A_Real-Time_Machine_Learning-Based_Public_Transport_

Bus-Passenger_Information_System/links/61bbdfcaa6251b553ac3d709/A - Real - Time - Machine - Learning - Based - Public - Transport - Bus - Passenger-Information-System.pdf.

- [20] E. Barsoum, C. Zhang, C. Canton Ferrer y Z. Zhang, “Training Deep Networks for Facial Expression Recognition with Crowd-Sourced Label Distribution,” en *ACM International Conference on Multimodal Interaction (ICMI)*, 2016.
- [21] microsoft, *GitHub - microsoft/FERPlus: This is the FER+ new label annotations for the Emotion FER dataset*. [En línea]. Disponible en: <https://github.com/Microsoft/FERPlus>.
- [22] *Challenges in Representation Learning: Facial Expression Recognition Challenge*. [En línea]. Disponible en: <https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data>.
- [23] *CK+ Dataset*. [En línea]. Disponible en: <https://paperswithcode.com/dataset/ck>.
- [24] M. Lyons, M. Kamachi y J. Gyoba, *The Japanese Female Facial Expression (JAFFE) Dataset*, abr. de 1998. [En línea]. Disponible en: <https://zenodo.org/records/3451524>.
- [25] *Welcome to Python.org*. [En línea]. Disponible en: <https://www.python.org/>.
- [26] K. Team, *Keras documentation: Keras 3 API documentation*. [En línea]. Disponible en: <https://keras.io/api/>.
- [27] *NumPy Website*. [En línea]. Disponible en: <https://numpy.org/>.
- [28] *Pandas documentation: Python Data Analysis Library*. [En línea]. Disponible en: <https://pandas.pydata.org/>.
- [29] *Matplotlib — Visualization with Python*. [En línea]. Disponible en: <https://matplotlib.org/>.
- [30] *Seaborn: statistical data visualization — seaborn 0.13.2 documentation*. [En línea]. Disponible en: <https://seaborn.pydata.org/>.
- [31] *Employment and Investigation of Various CNN Models and Datasets for Facial Expression Recognition and Classification*. [En línea]. Disponible en: https://ieeexplore.ieee.org/abstract/document/10415947?casa_token=Y5g50gMG9RYAAAAAxK02AVCySEMRqe4QI92gRX9_ED3E3F5DaajheAeEHkUBNFFq5P6m0Q1jQnp
- [32] O. C. Oguine, K. J. Oguine, H. I. Bisallah y D. Ofuani, *Hybrid Facial Expression Recognition (FER2013) Model for Real-Time Emotion Classification and Prediction*, jun. de 2022. [En línea]. Disponible en: <https://arxiv.org/abs/2206.09509>.
- [33] *BOE-A-2018-16673 Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales*. [En línea]. Disponible en: <https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673>.

- [34] *Reglamento general de protección de datos (RGPD)*. [En línea]. Disponible en: <https://eur-lex.europa.eu/ES/legal-content/summary/general-data-protection-regulation-gdpr.html>.
- [35] *BOE-A-2012-5043 Real Decreto 662/2012, de 13 de abril, por el que se establece el marco para la implantación de los sistemas inteligentes de transporte (SIT) en el sector del transporte por carretera y para las interfaces con otros modos de transporte*. [En línea]. Disponible en: <https://www.boe.es/eli/es/rd/2012/04/13/662>.
- [36] *Ley de IA de la UE: primera normativa sobre inteligencia artificial*. [En línea]. Disponible en: <https://www.europarl.europa.eu/topics/es/article/20230601ST093804/ley-de-ia-de-la-ue-primer-normativa-sobre-inteligencia-artificial>.
- [37] L. Carnaúba, *Guía Salarial IT 2024*, dic. de 2022. [En línea]. Disponible en: <https://ytalentfy.com/guia-salarial-it-2024/>.