

Visión Artificial

Tema 15. Aplicaciones actuales del tratamiento de la señal

Índice

Esquema

Ideas clave

15.1. ¿Cómo estudiar este tema?

15.2. Biometrics

15.3. Self-driving car

15.4. Retos y limitaciones de la percepción computacional

A fondo

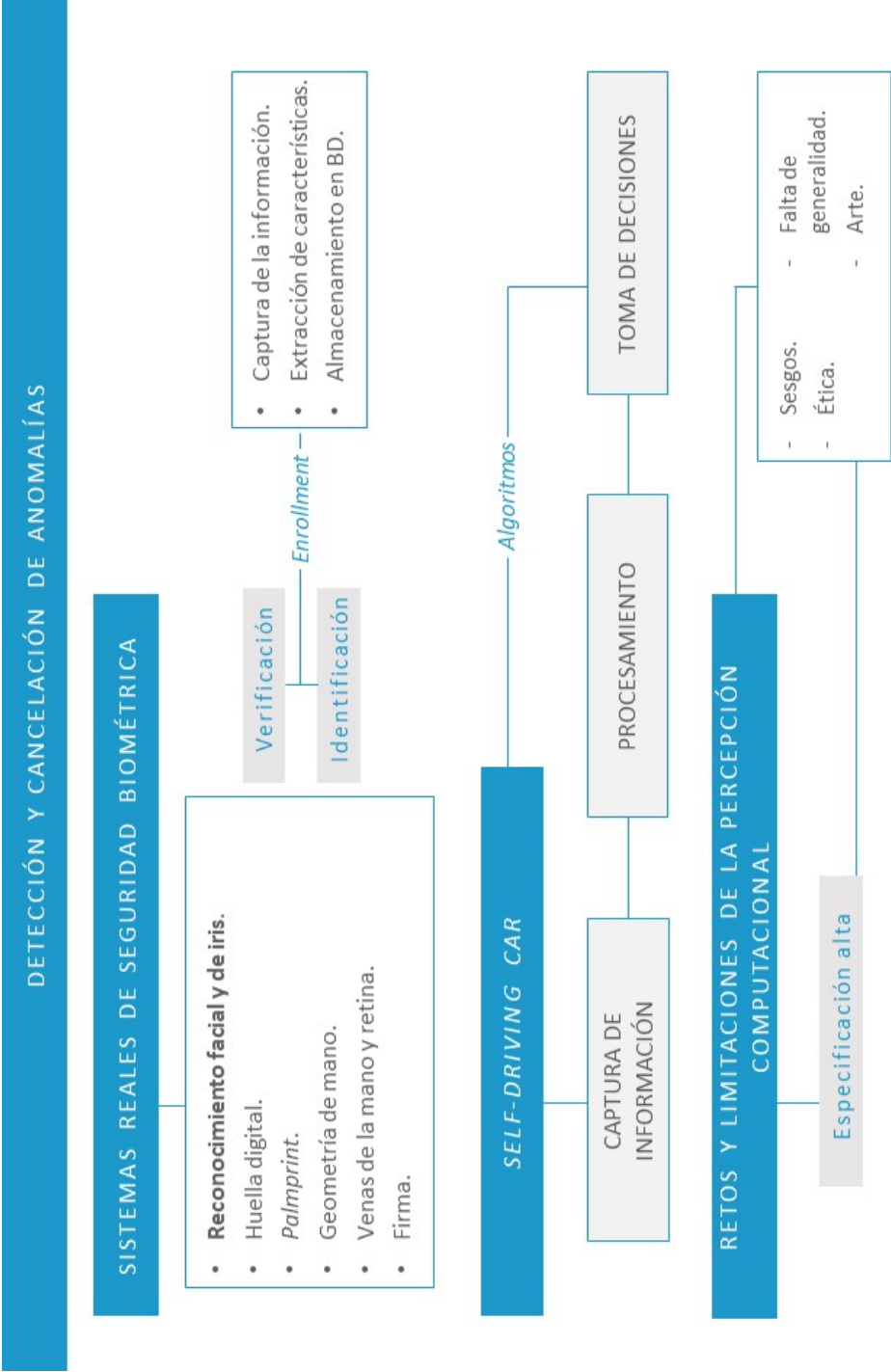
Biometric Devices

Reconocimiento del iris

Deep Learning and self-driving cars

Bibliografía

Test



15.1. ¿Cómo estudiar este tema?

Para estudiar este tema deberás leer con atención las ideas clave que se desarrollan a continuación.

Como último tema de la asignatura, la idea es aplicar los conocimientos presentados a dos tecnologías actuales y entender qué algoritmos y soluciones son los más adecuados para cada problema. Además, se presenta un conjunto de retos y limitaciones que son el trabajo actual y futuro de las investigaciones asociadas a percepción computacional.

15.2. Biometrics

La biometría es el conjunto de técnicas de reconocimiento de patrones capaces de verificar e identificar a una persona de forma unívoca. Actualmente, la seguridad y protección de la información reside en tres pilares:

- ▶ «Lo que sabes», también conocido como **contraseñas** en todas sus modalidades: usuario y contraseña corriente, sistemas de clave pública y privada, etc.
- ▶ «Lo que tienes», también conocido como **token**, consistente en identificar o más bien validar a un usuario por tener un determinado objeto/texto con información que únicamente el usuario posee. Este *token* es comúnmente usado en los bancos de dos maneras:
 - Mediante el envío de un mensaje al teléfono móvil, siendo en este caso el móvil el propio *token*.
 - Tarjeta de coordenadas, donde el banco te indica un número de dos cifras y el usuario debe completar con 3-4 cifras las correspondientes a dichas dos cifras.
- ▶ «Lo que eres», la *biometría*: rasgos únicos extraídos de propiedades físicas del individuo como puede ser el iris, la retina o la propia huella digital.

La seguridad de la biometría reside en «lo que eres»: en características físicas y de comportamiento únicas para cada usuario o que, sin ser tan únicas, permiten al menos poder identificar a los usuarios lo suficiente como para distinguirlos del resto de usuarios. Es importante distinguir dos usos dentro de estos sistemas biométricos:

Verificación. En la verificación, el **usuario proporciona algún tipo de identificación** y el sistema biométrico comprueba que realmente es quien dice ser. Para ello, el sistema biométrico compara los rasgos extraídos con el patrón (también llamado *template*) biométrico de dicho usuario. Si dicha comparación no supera un umbral, entonces el usuario es rechazado. Por el contrario, si dicha comparación es satisfactoria, el sistema verificaría al individuo positivamente.

Identificación. Al contrario que en la verificación, el usuario no proporciona **ningún tipo de identificación previa**, luego el sistema tiene que reconocer/identificar de qué usuario se trata. Esto lo realiza comparando los rasgos biométricos adquiridos en el sistema de captura con todos y cada uno de los patrones (*templates*) almacenados en la base de datos.

En cualquier caso, parece obvio que el usuario ha de estar primeramente registrado en el sistema. Este proceso se conoce como *enrollment*. Y para ello se realizan las siguientes operaciones:

- ▶ Captura de la información biométrica mediante sensores especializados: cámaras de infrarrojo, sensores de huella, etc.
- ▶ Extracción de características para la creación de un patrón.
- ▶ Almacenamiento en la base de datos.

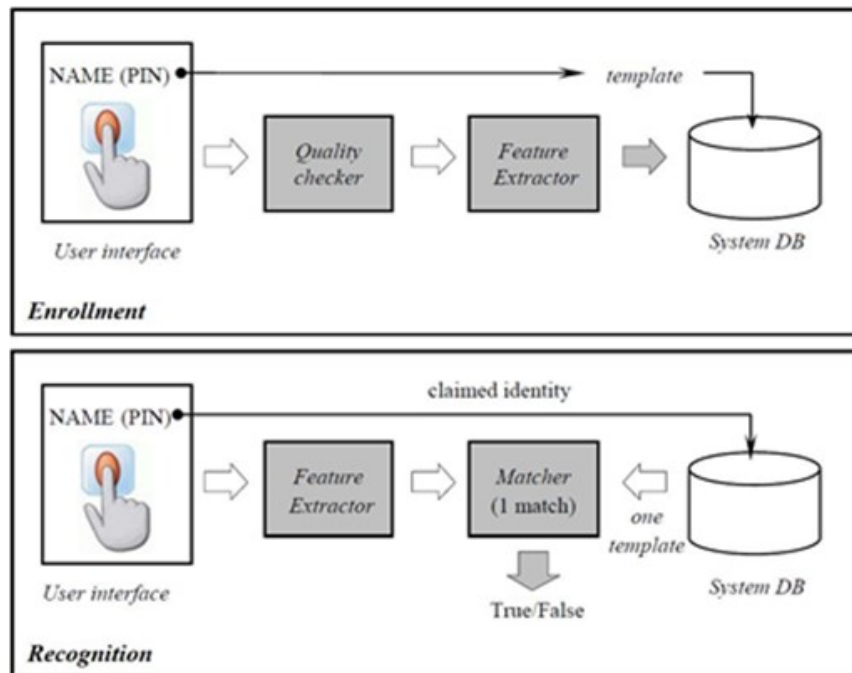


Figura 1. Funcionamiento de un sistema biométrico en sus dos fases más típicas. Fuente:

http://biometrics.cse.msu.edu/pub/secure_biometrics.html

En la imagen anterior vemos esquemáticamente cómo son las fases de un sistema biométrico: dar de alta un usuario (*enrollment*) y comprobación del usuario (*recognition*); en este caso concreto, la comprobación se trata de una verificación.

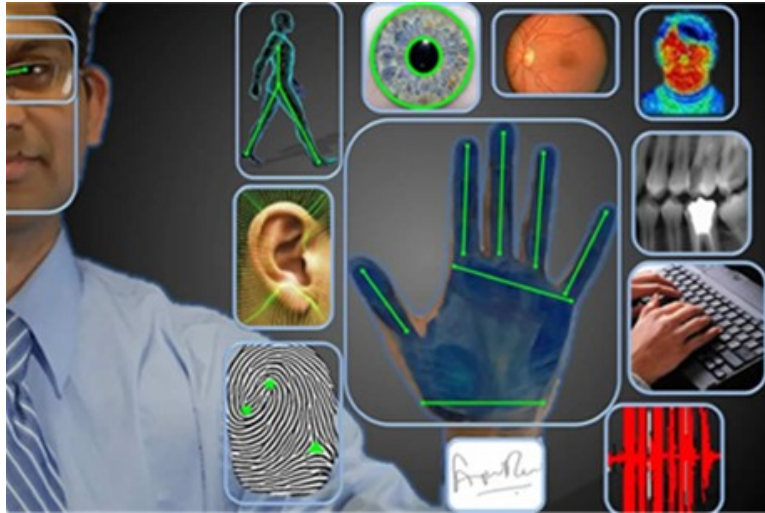


Figura 2. Rasgos físicos más comunes empleados en el reconocimiento biométrico. Fuente:

<https://www.vrstech.com/biometric-systems-security-services.html>

A continuación, presentamos un resumen de las técnicas más comúnmente empleadas en biometría: reconocimiento facial y biometría de iris, junto con los algoritmos tanto de segmentación como de extracción de características más extendidos en la literatura para cada técnica.

Reconocimiento facial

La cara, desde un punto de vista geométrico, ofrece características muy identificativas: la distancia entre ojos, proporción de la nariz, etc.

Para la **segmentación**, un detector de color de piel (segmentación basada en color) suele ser suficiente, aunque existen enfoques donde se persigue la segmentación basada con infrarrojos o bien con sensores de temperatura, mucho más avanzados y costos económicamente.

Para la **extracción de características**, el concepto de *eigenfaces* es el más extendido. Este método emplea el uso de *Principal Component Analysis* para encontrar las componentes principales en una imagen; realiza, además, una descomposición en vectores propios (de ahí el nombre de *eigenfaces*) que proporciona un resultado muy preciso con una computación razonable.

Las ventajas de la biometría basada en reconocimiento facial se basan en su poca invasividad: el usuario no necesita colaborar sustancialmente para que el proceso de verificación/identificación se lleve a cabo. Las dificultades que presenta el reconocimiento facial son su dependencia de agentes externos como la iluminación o el uso de gafas, barba o incluso heridas que puedan modificar los rasgos biométricos.

En la siguiente imagen vemos que, a partir de diferentes caras del usuario, se obtienen sus rasgos más significativos.



Figura 3. Ejemplo visual de eigenfaces. Fuente: <https://www.emaze.com/@AQCLWRRF/Facial-Recognition>

Reconocimiento de iris

Uno de los más potentes a nivel de precisión por la gran entropía y unicidad que ofrece el iris del ojo humano. Cada ojo, incluso entre izquierdo y derecho, es único a nivel de textura.

La **segmentación** está basada en tres principios:

- ▶ Segmentación basada en **color**: aunque de las más sencillas, es la menos precisa de todas debido a su gran variabilidad de iluminación.
- ▶ Segmentación basada en **forma**: asumiendo que el iris es circular, intentar detectar zonas circulares dentro de la imagen del ojo. Tiene mucho sentido emplear algoritmos de crecimiento de regiones en este tipo de segmentaciones.
- ▶ Segmentación basada en **textura**: empleada en la mayoría de casos debido a que las capturas se suelen realizar con sensores infrarrojos.

La **extracción de características** viene dada por filtros de Gabor, aunque existen infinitas posibilidades, todos ellos basados en texturas: *Markov Random Fields*, Wavelets, DCT, etc.

La comparación de unas características biométricas basadas en iris con su correspondiente patrón biométrico suele ser muy sencillo computacionalmente (distancia euclídea, distancia de Hamming, etc.). De esta manera, la comparación es muy rápida y permite poder tomar decisiones rápidamente.

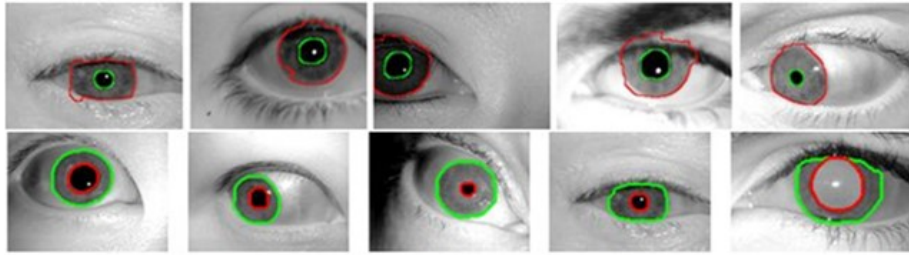


Figura 4. Ejemplos de segmentación de iris y pupila. Fuente:

<https://www.andrew.cmu.edu/user/thihoanl/Research.html>

Cuando se emplean cámaras infrarrojas, la segmentación basada en color no produce buenos resultados y hay que acudir a segmentaciones más avanzadas basadas en formas o en texturas.

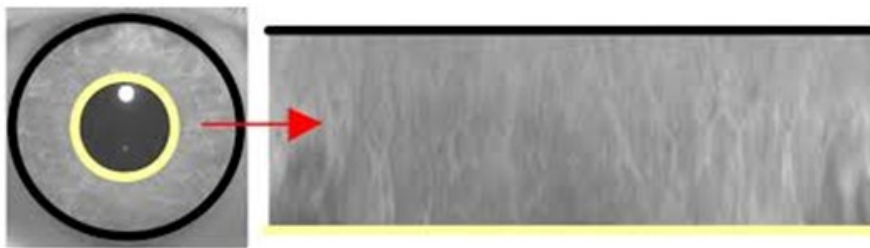


Figura 5. Transformación de coordenadas polares a cartesianas del tejido correspondiente al iris. Fuente:

<http://www.cs.princeton.edu/~andyz/irisrecognition>

Esta transformación facilita la posterior aplicación de filtros y comparación de patrones.

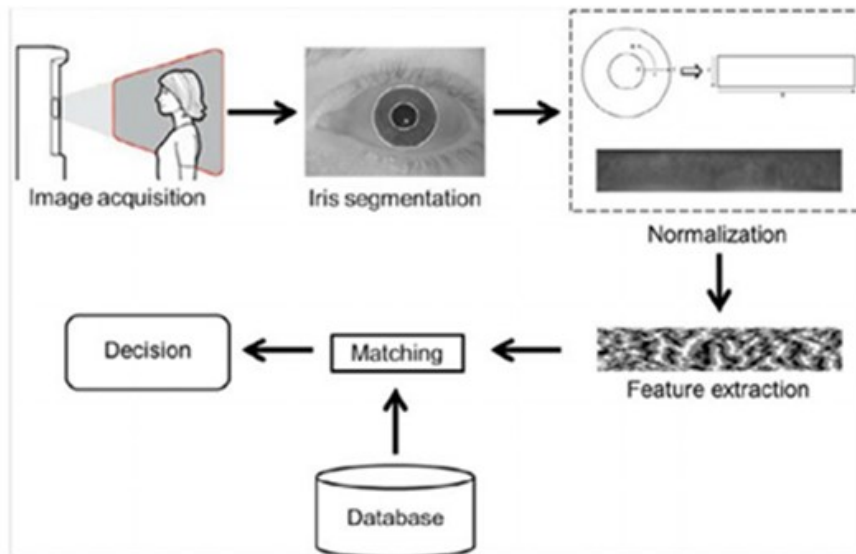


Figura 6. Esquema general de un sistema de reconocimiento biométrico basado en iris. Fuente:

<https://www.bayometric.com/biometric-iris-recognition-application/>

Y por último, se presenta un resumen breve de otras técnicas biométricas con los algoritmos de segmentación y extracción de características más comunes en la literatura.

Huella digital

Sin duda, una de las técnicas biométricas más empleadas, y no por ella la más sencilla.

- ▶ Sistemas de captura: suelen ser sistemas de infrarrojos o escáneres.
- ▶ Segmentación: muy sencilla, ya que se delega en el sensor toda la dificultad de la captura de la huella. Suele tratarse de una segmentación basada en clustering de dos clases junto con morfología matemática para eliminar errores.

- ▶ Extracción de características: propiedades geométricas de los patrones de las líneas de la huella. Dichas propiedades son muy similares a los algoritmos SIFT.

Palmprint

Esta técnica consiste en leer las líneas de la palma de la mano, no solo las más visibles, también la textura; similar a la huella del dedo, pero en la palma de la mano.

- ▶ Sistemas de captura: muy variados, aunque los más usados son los que emplean fotografías infrarrojas.
- ▶ Segmentación: más compleja que en la huella; en este caso hay que identificar correctamente la palma de la mano y hacer una segmentación basada en intensidad.
- ▶ Extracción de características muy similares a la de la huella digital.

Geometría de mano

Esta técnica permite identificar a las personas en función de la forma de su mano.

- ▶ Sistemas de captura: muy sencillos, bastaría con una cámara del móvil.
- ▶ Segmentación: cuanto más sencillo sea el dispositivo de captura, más compleja será la segmentación. En este caso, la segmentación suele conllevar agregación multiescala, segmentación basada en color o crecimiento de regiones.
- ▶ La extracción de características está orientada a medidas relativas de los dedos, mano y muñeca.

Venas de la mano y retina

- ▶ Sistema de captura: infrarrojo obligatoriamente para poder ver a través de la piel.
- ▶ Segmentación: basada en color y en intensidades.
- ▶ Extracción de características: suele usarse filtros de Gabor y algoritmos que simulan redes comparando las venas a redes de nodos.

Firma

- ▶ Sistema de captura muy especializado: o bien un sensor con un bolígrafo específico (*online*), o bien una fotografía de una firma (*offline*).
- ▶ En el caso online no hay segmentación. En el *offline*, la segmentación es muy sencilla: binarización de la imagen.
- ▶ Extracción de características: se suelen utilizar métodos basados en parametrización de curvas y propiedades parecidas a SIFT. No obstante, la comparación en esta técnica requiere especial atención. Algoritmos basados en programación dinámica como Dynamic Time Warping suelen ofrecer resultados excelentes ya que trabajan muy bien con señales que pueden estirarse en el tiempo y en el espacio.

En definitiva, la biometría es uno de los campos, junto con las imágenes biomédicas, donde más se ha avanzado en la percepción computacional y donde más se está mejorando las capacidades humanas de reconocimiento.

15.3. Self-driving car

Recientemente están comenzando a surgir diferentes empresas capaces de crear coches que conducen por sí mismos. Muchas de estas marcas afirman que los coches son capaces de detectar más información (y procesarla) de la que un humano es capaz de detectar y procesar.

En este capítulo vamos a tomar el esquema inicial presentado en el tema de *Elementos de un sistema de percepción* para modelar la percepción computacional: captura de información, procesamiento y toma de decisiones.

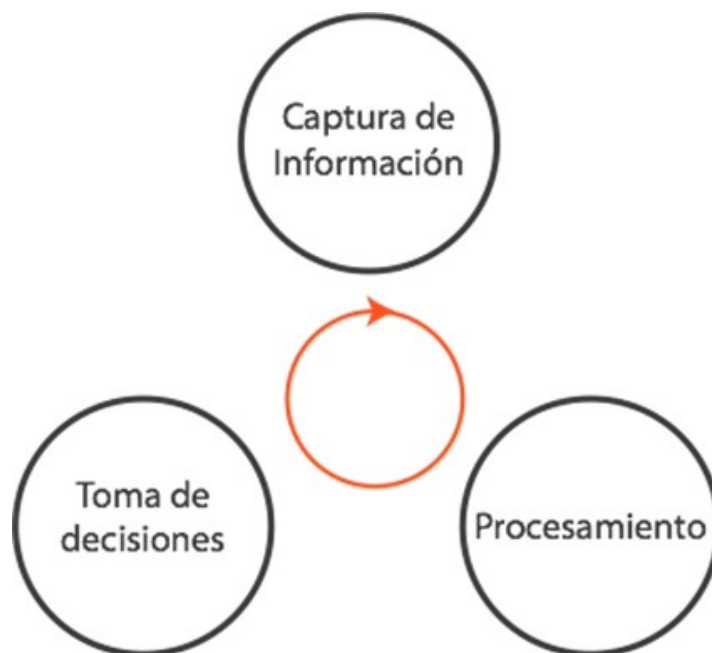


Figura 7. Esquema básico de percepción computacional.

Con respecto a la captura de información, existe una gran diversidad de sensores que en efecto superan a los humanos: infrarrojos, radares, sensores de aceleración, de movimiento, de humedad, posicionamiento, etc. Hoy en día no se contempla, pero posiblemente, en un futuro, la existencia de drones aéreos vigilando y proporcionando una vista aérea permita tener un control del tráfico enorme y poder

proporcionar a los *self-driving car* la posibilidad de entender todo el contexto de la situación.

A continuación, se presenta un breve esquema de todas las posibilidades sensoriales que ofrece un coche autodruido.

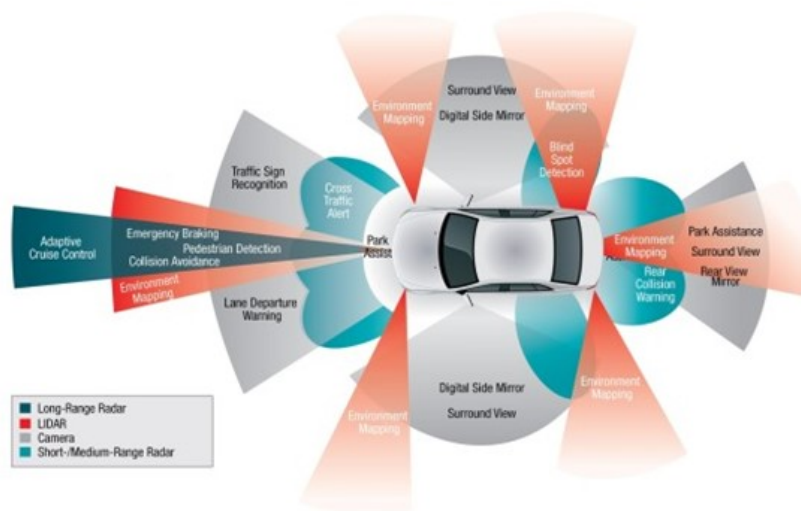


Figura 8. Descripción de sensores incluido en un self-driving car. Fuente:

<https://www.shofior.com/wordpress/self-driving-car-how-it-works/?lang=en>

Es obvio que para capturar toda esta información, sincronizarla, preprocesarla y limpiarla, entre otras tareas, hace falta un computador a bordo. De hecho, dicho computador está especialmente dedicado a estas tareas, considerando que el procesamiento de vídeo es una de las tareas computacionales más costosas. Además, en muchos casos los sensores son capaces de construir representaciones en 3D de la realidad, haciendo que los algoritmos de segmentación, como puede ser el algoritmo de morfología matemática del que hemos hablado en esta asignatura, aumenten sustancialmente en coste computacional, ya que el procesamiento ha de hacerse en el momento, en tiempo real.

De hecho, la computadora encargada del procesamiento puede ocupar el maletero del propio coche. Sin embargo, el reto está precisamente en el procesamiento de la información capturada por múltiples motivos.

- ▶ En primer lugar, se está capturando información de diversa naturaleza:
 - Información **unidimensional** como puede ser la información proveniente de los sensores.
 - Información **bidimensional** como las propias imágenes o vídeos que se capturan para analizar el movimiento.
 - Información **tridimensional** como son los radares que dotan al coche de la posibilidad de entender la profundidad en el espacio.
- ▶ En segundo lugar, la información ha de procesarse a la misma velocidad: no puede procesarse primero la sensorización de humedad (por poner un ejemplo) y posteriormente el vídeo procedente de la cámara. Un desalineamiento temporal en este caso produciría que la decisión final (de la cual hablaremos más adelante) sea errónea y se esté tomando con base en datos completamente incorrectos.
- Además, dicho procesamiento ha de contemplar la posibilidad de que no pueda devolver ningún resultado satisfactorio, con lo que precisa de una lógica necesaria para actuar en estos casos: una mala segmentación, cambio drástico de luz, etc.
- Por último, debe almacenar todos los datos procesados, ya que moralmente y en caso de accidente se deberá analizar qué sucedió, qué vio el coche y por qué actuó de esa manera.



Figura 9. Hardware necesario para procesar toda la información capturada por un self-driving car. Fuente: <https://www.nytimes.com/es/2016/02/10/los-autos-que-se-manejan-solos-todavia-necesitan-a-los-humanos/>

¿Qué algoritmos podría aplicar un *self-driving car*?

Algoritmos de **segmentación** con la finalidad de:

- ▶ Separar el fondo del vídeo que se captura eliminando, por ejemplo, las zonas de césped, o detectando qué zona se corresponde con un cambio de rasante.
- ▶ Detección de diferentes autos dentro de la vía.

Algoritmos de **detección** de bordes con la finalidad de:

- ▶ Detectar correctamente las fronteras de la vía.
- ▶ Detectar la existencia de líneas en la carretera, como puede ser la línea continua.

Algoritmos de **extracción de características** con la finalidad de:

- ▶ Detectar objetos en movimiento; mediante Wavelets y filtros de Gabor se pueden detectar texturas diferentes a las del asfalto o cambios de texturas en el paisaje.
- ▶ Detección de señales; mediante el algoritmo SIFT se puede identificar señales de tráfico fácilmente.

Algoritmos de **corrección de ruido**: un ejemplo claro son las señales de tráfico. En caso de lluvia o condiciones adversas, puede que el sensor de señales confunda la señal de prohibido; ir a más de 60 km/h con la de ir a 80 km/h, con las correspondientes consecuencias.

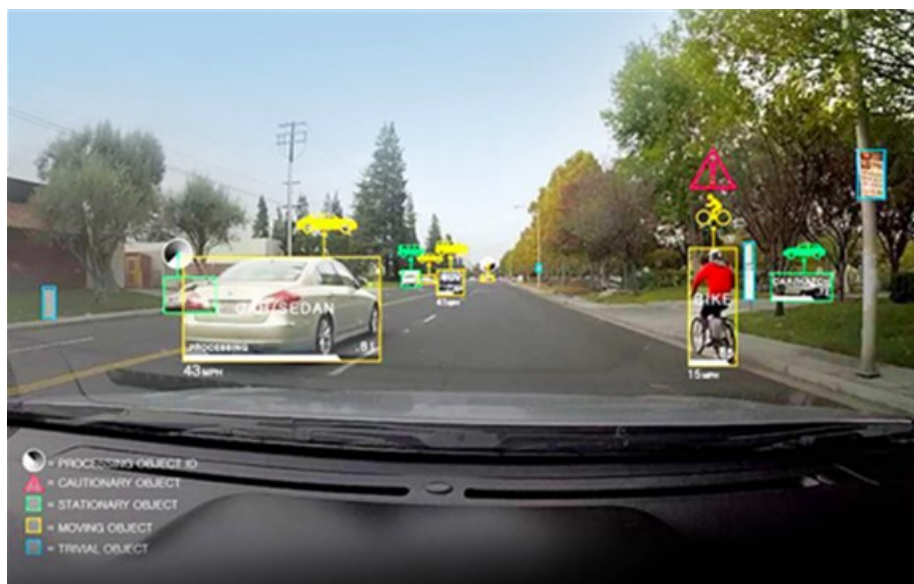


Figura 10. Detección de objetos en un *self-driving car*. Fuente:

<http://www.eenewsautomotive.com/news/nvidia-computer-processing-hub-self-driving-cars>

La identificación de diferentes objetos, así como el reconocimiento y clasificación de los mismos favorece la toma de decisiones.



Figura 11. Segmentación de imágenes aplicada a un *self-driving car*. Fuente:

<https://blogs.nvidia.com/blog/2016/01/05/eyes-on-the-road-how-autonomous-cars-understand-what-theyre-seeing/>

En diferentes colores se aprecian los diferentes tipos de objetos en la imagen: asfalto (morado), personas (verde), señales de tráfico (amarillo) y vehículos (rojo).

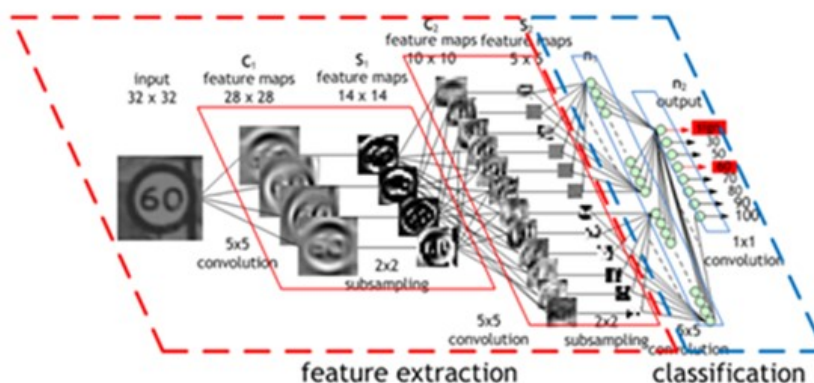


Figura 12. Ejemplo de red neuronal para la extracción de características y clasificación de una señal de tráfico. Fuente: <https://towardsdatascience.com/beginning-my-journey-in-self-driving-car-udacity-nano-degree-a39d898658a2>

En esta imagen se aprecia cómo, tras la clasificación, la imagen se detecta correctamente como una señal de limitación de 60 km/h.

Y por último, el *self-driving car* consta del **módulo de toma de decisiones**. En este caso, la ética aún no se ha pronunciado sobre qué decisiones debe tomar un *self-driving car*. Por ahora, lo único que hace en una situación anómala es avisar al conductor para que tome el control del vehículo.

Pero en el caso real de un conductor distraído o de un coche completamente autónomo, ¿qué decisiones podrá tomar un vehículo? ¿Quién sería el responsable si el vehículo toma una decisión incorrecta? ¿Y si, en el caso extremo, tiene que elegir entre salvar la vida de sus ocupantes o la de los viandantes? Y en un caso más liviano, una multa de tráfico, ¿de quién sería la responsabilidad?

Como todo módulo de percepción computacional proporciona un soporte a la decisión del humano que, en función de los datos proporcionados por el sistema, decidirá qué opción es la más adecuada.

15.4. Retos y limitaciones de la percepción computacional

En la percepción computacional existe aún mucho espacio de mejora. Los algoritmos que hemos visto, por su gran generalidad, requieren de una **especificación muy alta** que hace que pierdan esa capacidad de generalidad.

De hecho, algoritmos que detecten una mano en una imagen, por ejemplo, no podrán usarse si se cambia el sistema o las condiciones de captura. En la mayoría de los casos, esto impide que los sistemas de percepción computacional sean generales.

La incorporación de nuevas técnicas de aprendizaje como Deep Learning está haciendo que esa meta esté más cerca, pero aún queda mucho camino por recorrer.

Nos gustaría poner un ejemplo llamativo de hasta qué punto se está exigiendo a la inteligencia artificial el ser capaz de tomar decisiones o de decidir una clasificación para una imagen.

Existen competiciones donde se ponen a prueba los algoritmos de clasificación de imágenes. Uno de ellos es la competición *Muffin or Chihuahua*. El parecido entre imágenes de ambas clases es comprensible y confunde enormemente al clasificador y al sistema de percepción computacional que hay detrás.



Figura 13. Competición Muffin or Chihuahua. Fuente: <https://medium.freecodecamp.org/chihuahua-or-muffin-my-search-for-the-best-computer-vision-api-cbda4d6b425d>

No obstante, ¿qué se puede hacer para mejorarlo?

- ▶ Incluir información relevante a la textura con la intención de distinguir el muffin del pelo del perro.
- ▶ No usar algoritmos como SIFT ya que podrían confundir la mayoría de estas imágenes.
- ▶ Aplicar algoritmos basados en Wavelets para poder potenciar los detalles.
- ▶ Aplicar características que modele la simetría de un perro en comparación a la asimetría (o simetría casual) de un *muffin*.

En cualquier caso, el sistema de percepción resultante sería muy específico y no podría usarse para identificar ningún otro tipo de objeto o hacer ningún otro tipo de comparación.

Por último, indicamos una lista de los retos a los que se enfrentan actualmente estos sistemas de percepción computacional. Esta lista es muy dinámica y seguramente, en la fecha en la que escribimos estas limitaciones, ya han surgido enfoques o soluciones que afrontan dichas limitaciones porque un aspecto positivo de la percepción computacional es su amplia comunidad y alto interés.

Sesgos. Uno de los aspectos más inquietantes es el propio sesgo humano: ¿hasta qué punto puede ser que un algoritmo aprenda o imite exactamente cómo decide un humano y no desarrolle toda la potencialidad?

Ética. ¿Se puede confiar en lo que una máquina decide por sí misma? ¿Se puede confiar la vida de personas a un sistema de conducción automática? ¿Y a una vigilancia automática? ¿Podremos vivir con sistemas de percepción integrados en nuestros sistemas de visión humanos?

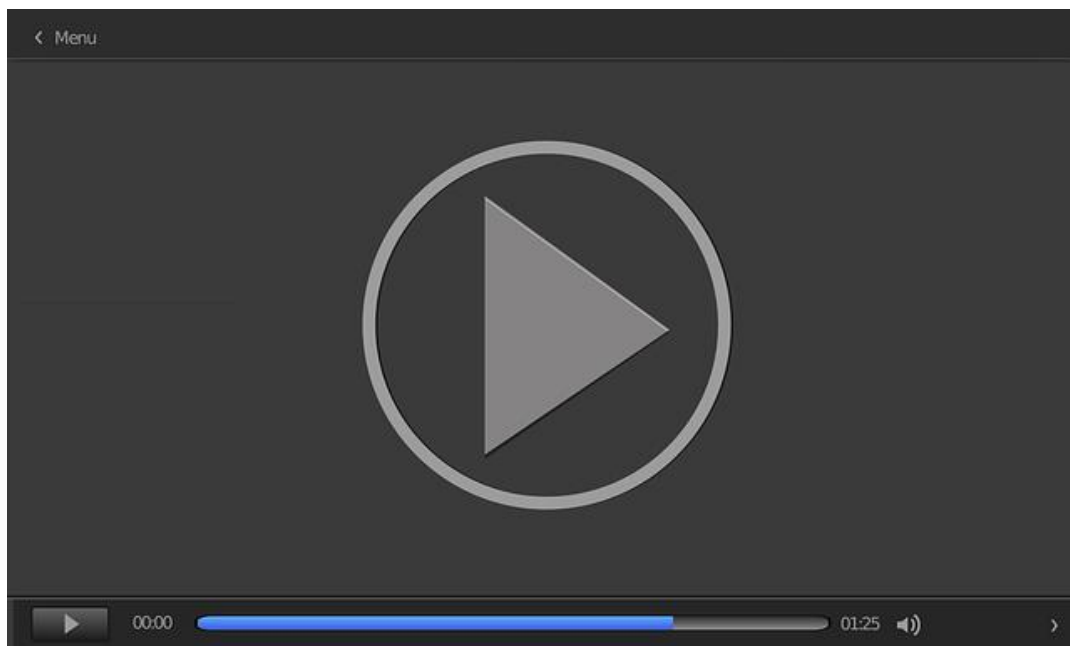
Falta de generalidad. ¿Se podrá crear un algoritmo capaz de aprender y entender qué hay en la imagen sin necesidad de que dentro lleve múltiples clasificadores? ¿Podrá la percepción computacional deducir nuevos conceptos de una imagen así como lo hace un humano?

Arte. Puesto que estos sistemas pueden percibir imágenes, ¿podrían crear las suyas propias? Y en ese caso, ¿serían completamente originales?

Biometric Devices

Star Link Communication Pvt. Ltd. (2016, septiembre 30). *How Fingerprint Recognition Works? // Biometric Devices // Star Link* [Vídeo]. YouTube. <https://www.youtube.com/watch?v=AZkc48X5yck>

Vídeo que explica de forma sencilla y básica cómo funciona la lectura de la huella digital en dispositivos biométricos y el reconocimiento de imagen.



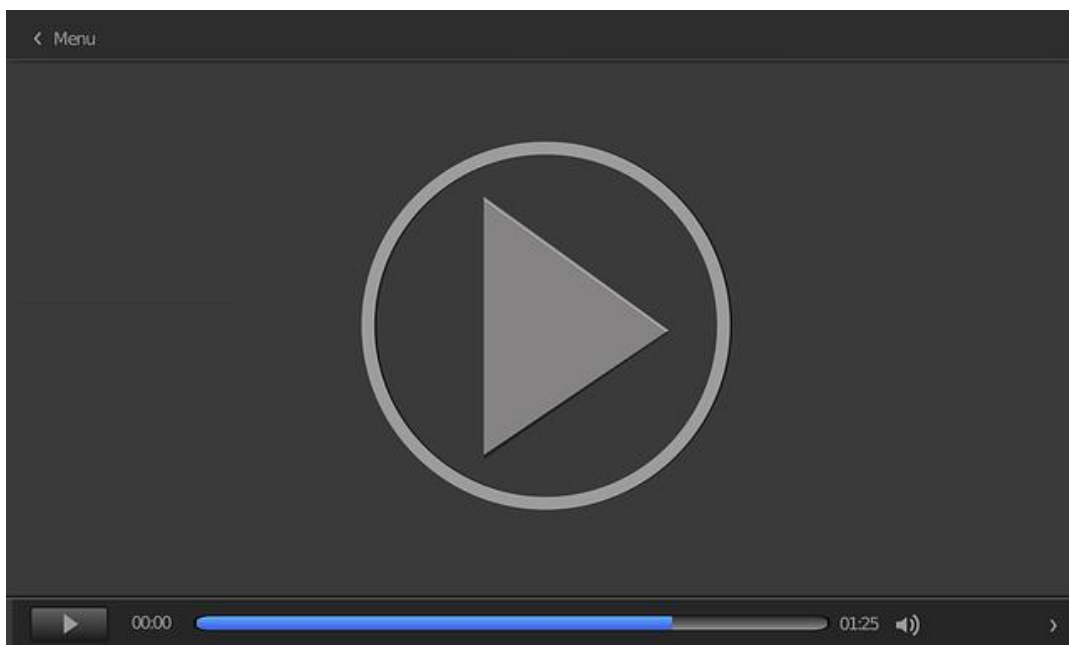
Accede al vídeo:

<https://www.youtube.com/embed/AZkc48X5yck>

Reconocimiento del iris

SAIConference. (2015, noviembre 12). *Iris Recognition by Prof. John Daugman* [Vídeo]. YouTube. <https://www.youtube.com/watch?v=KyDoFrojEYk>

Conferencia de John Daugman, creador de la biometría de iris, en la que explica cómo funciona el algoritmo.



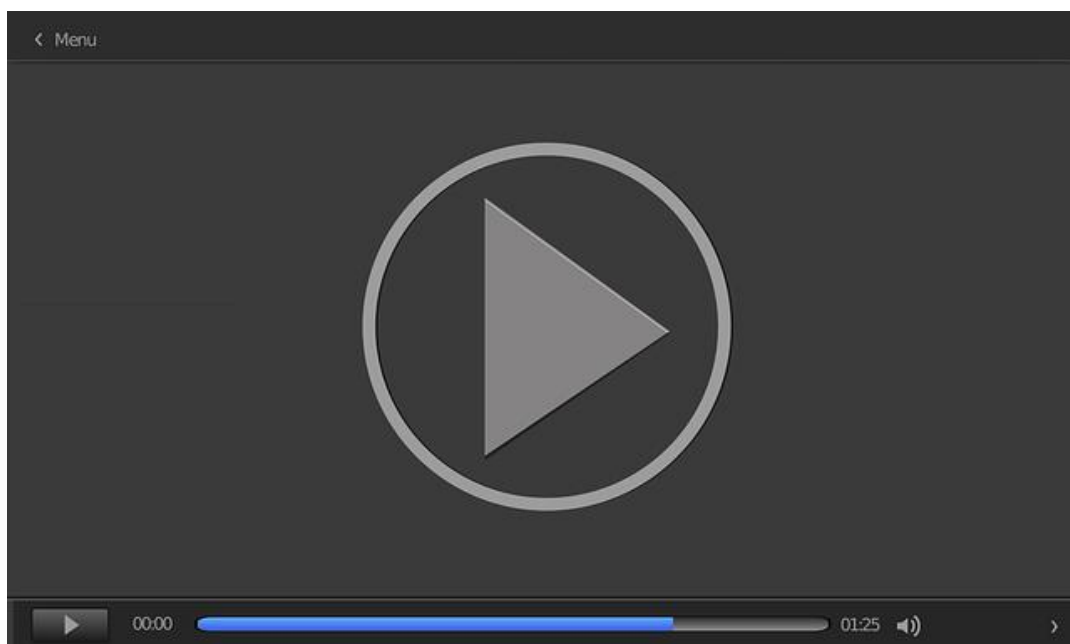
Accede al vídeo:

<https://www.youtube.com/embed/KyDoFrojEYk>

Deep Learning and self-driving cars

Lex Fridman. (2017, enero 16). *MIT 6.S094: Introduction to Deep Learning and Self-Driving Cars* [Vídeo]. YouTube. <https://www.youtube.com/watch?v=1L0TKZQcUtA>

Conferencia de Lex Fridman en el MIT sobre *Deep Learning* y *self-driving cars*.



Accede al vídeo:

<https://www.youtube.com/embed/1L0TKZQcUtA>

Bibliografía

González, R. C. y Woods, R. E. (2008). *Digital image processing*. New Jersey: Pearson Education.

1. Los actuales métodos de *self-driving car* realizan la conducción automática sin hacer segmentación puesto que es computacionalmente muy costosa. Sin embargo, realizan una extracción de características muy minuciosa:
 - A. Falso
 - B. Verdadero.

2. Un sistema de conducción autónomo, ¿podría realizar el análisis de los datos en la nube? Es decir, en un servidor remoto enviando los datos a través de la red móvil:
 - A. Verdadero, siempre y cuando la conectividad sea constante y de un alto ancho de banda.
 - B. Falso, en ningún caso.
 - C. Falso, tiene que realizarse dentro del coche para asegurar tiempos de respuesta (rapidez de procesamiento) muy bajos.

3. Los sistemas autónomos de conducción tienen entre sus limitaciones:
 - A. Dificultad en la toma de decisiones complejas (como accidentes).
 - B. Los sistemas de captura pueden sufrir alteraciones debidas al entorno (lluvia, nieve, etc.).
 - C. Necesitan de un hardware adicional para realizar el cómputo del análisis.
 - D. Todas las anteriores son correctas.
 - E. Solo A y B son correctas.

4. ¿Qué ventaja tendría combinar un sistema de conducción automática con biometría?

- A. La detección del conductor y asegurar que solo el conductor (o conductores) pueden montarse en él.
- B. Reconocer caras durante la conducción y registrar quién está pasando por qué sitio.
- C. Reconocer el comportamiento fisiológico (constantes vitales, tasa cardiaca, respiración, etc.) y ver si está nervioso el conductor.
- D. Todas las anteriores.

5. La biometría estudia el reconocimiento de individuos en función de patrones físicos o de comportamiento:

- A. Falso.
- B. Verdadero.

6. La biometría basada en huella precisa de algoritmos de segmentación complejos:

- A. Falso
- B. Verdadero.

7. La biometría basada en iris necesita algoritmos de segmentación basada en color para identificar la pupila:

- A. Si es solo la detección de pupila, la segmentación basada en color puede ser un buen enfoque.
- B. Falso. Aunque la pupila sea negra, necesita obligatoriamente de una segmentación basada en regiones.
- C. La biometría de iris no necesita identificar la pupila.

8. ¿Qué técnica puede ser usada como reconocimiento biométrico?
 - A. La manera de caminar.
 - B. La manera en la que tecleamos en un ordenador.
 - C. La forma de la oreja.
 - D. Todas ellas pueden usarse como reconocimiento biométrico.

9. Señala la mayor limitación de los sistemas de percepción computacional:
 - A. Falta de generalización: los sistemas actualmente resuelven problemas muy concretos y por ello son sistemas muy especializados.
 - B. Son algoritmos lentos y computacionalmente costosos y por ello no pueden ejecutarse en dispositivos móviles.
 - C. Pueden tomar decisiones sesgadas.
 - D. Todas las anteriores.

10. Si un sistema de percepción computacional no sabe distinguir dos clases de objetos, ¿qué solución es la más óptima?
 - A. Seguro que cambiar el algoritmo de segmentación ayuda a mejorar la distinción.
 - B. Cambiar la extracción de características a tipos de extracciones donde las diferencias, visualmente, sean más obvias (cambiar de formas a texturas, por ejemplo).
 - C. Cambiar el clasificador que toma la decisión.
 - D. Ninguna de las anteriores.