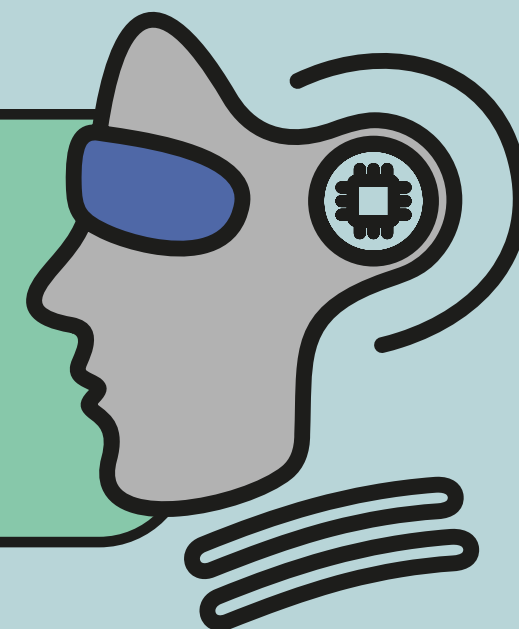


AUTENTICACIÓN FACIAL

UN HUMANOIDE NOS PUEDE RECONOCER

Aida A. Aparicio-Arroyo*, Iván Olmos-Pineda*, J. Arturo Olvera-López*

* Benemérita Universidad Autónoma de Puebla, Puebla, México.
Contacto: aparicio.200806025@gmail.com



EL AVANCE TECNOLÓGICO, UN PEQUEÑO VISTAZO

Gracias al avance tecnológico se han desarrollado dispositivos que sirven de apoyo en diferentes tareas cotidianas: teléfonos celulares, computadoras de escritorio y portátiles (herramientas de cálculo de propósito general), automóviles, e incluso diferentes tipos de robots (brazos articulados, exoesqueletos), entre otros. En el futuro, estos últimos tendrán un mayor grado de influencia, extendiendo

las aplicaciones con las que cuentan en la actualidad, no sólo en la industria (ensamble de diferentes tipos de objetos), o dentro del ámbito militar (búsqueda de explosivos), sino incluso en el área de la medicina (cirugías de alto riesgo). Un ejemplo de esto es el desarrollo de una prótesis robótica para quienes hayan perdido extremidades por algún accidente, la cual ejecuta de manera automática o a

través de estímulos biológicos o neuronales.

Incluso, se estima que en algunas décadas existan androides capaces de realizar labores de servicio y acompañamiento, así como en múltiples ocasiones se han mostrado en películas de ciencia ficción como *Yo robot*, *El hombre bicentenario*, *RoboCop* o *A.I. Inteligencia artificial*. En este sentido, es de especial interés para la robótica diseñar mecanismos capaces de interactuar de forma natural con los humanos como si se tratase de seres iguales.

Con base en esta idea, un robot puede llegar a tener un aspecto físico muy parecido al de un humano, e incluso puede imitar las expresiones faciales, se le conoce como humanoide (RH). Como el desarrollado por Wood et al. (2021), que sirve de ayuda en terapias de niños con autismo. En la figura 1 se muestra la imagen de un RH, diseñado por la empresa Hanson Robotics, donde se pueden

observar los avances en la construcción de aparatos parecidos a un humano.

Los RH se diseñan a partir de un conjunto de circuitos electrónicos que permiten controlar servomotores que activan el movimiento de diferentes componentes, como la cabeza, brazos y piernas. En estas animaciones se incluyen, en algunos casos, las expresiones fisonómicas.

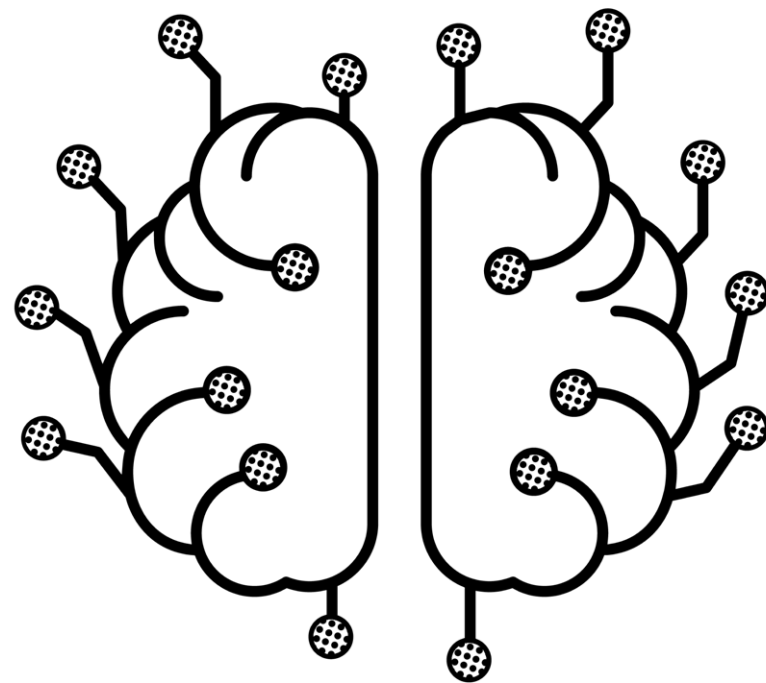
En la figura 2 se muestra la electrónica de un RH cuyos comandos de control son recibidos mediante una computadora externa. El ejemplo ilustrado tiene la capacidad de mover la cabeza, así como de generar expresiones faciales (movimientos de ojos, cejas, boca). Aunado a lo anterior, éstos suelen equiparse con sensores, micrófonos y cámaras de video que permiten adquirir datos del entorno. Con lo anterior se busca hacerlos más realistas, pues el objetivo es que puedan interactuar con los humanos de forma natural.



Figura 1. Arthur, robot humanoide propiedad de la BUAP, México.



Figura 2. Un humanoide y sus componentes electrónicos.



INTELIGENCIA ARTIFICIAL, EL FUTURO

Para lograr la integración de los humanoides en la vida diaria, aún hay que resolver muchas tareas. Por mencionar alguna: si se quiere que entable un diálogo en lenguaje natural con un humano, el aparato debe tener visión artificial, para poder ubicar a la persona con quien entable el diálogo; capacidad auditiva, para reconocer los sonidos que emite su interlocutor; léxico, para poder estructurar y emanar sonidos que representan las frases que desea expresar,

e incluso capacidad de movimientos corporales, que son usados por los humanos para enfatizar o no la comunicación.

Tomando en cuenta estas cuestiones, y enfocándose en el objetivo, el área de investigación de visión por computadora se encarga de procesar y analizar imágenes obtenidas a través de cámaras fotográficas o de video, las cuales son procesadas mediante la aplicación de algoritmos computacionales

para lograr diferentes metas: detectar objetos de interés en una escena (personas, animales), estimar trayectorias de objetos (conducción autónoma de vehículos), clasificar objetos (diferenciar uno de otro en una escena), entre muchas otras funciones.

Considerando las diferentes metas de la visión computacional, algunas aplicaciones de la detección de objetos es el reconocimiento facial y, por ende, la autenticación de seres humanos.

¿RECONOCER O AUTENTICAR PERSONAS?

Como se mencionó anteriormente, una de las tareas involucradas para el desarrollo de los sistemas a implementarse en un humanoide es la visión por computadora. En ésta existen diferentes áreas de interés como reconocer y autenticar; aunque podrían interpretarse como un mismo concepto o proceso, son totalmente diferentes. Reconocer responde a la pregunta ¿quién es la persona?; autenticar, ¿es quién realmente dice ser? (Jain et al., 2011).

El proceso de autenticación consiste (figura 3), en

primer lugar, en la adquisición de un conjunto de imágenes de rostros de personas, de preferencia una cantidad numerosa de fotografías de la misma persona, con el objetivo de tener una buena muestra para el entrenamiento del clasificador; a este conjunto de cuadros se le realiza un preprocesamiento, con el fin de corregir algunas imperfecciones de calidad. Como segunda etapa se encuentra la extracción de los rasgos faciales de cada uno de los retratos de todos los sujetos de prueba, éstos son almacenados en

un vector particular, el cual puede contener desde tres hasta n características (en párrafos posteriores se comentará la importancia del número de particularidades a extraer).

Posteriormente, el vector es procesado en un clasificador para obtener un modelo de entrenamiento que será utilizado en la etapa final: la autenticación. En esta etapa, el modelo es probado con nuevas imágenes de rostros y el resultado será la autenticación del individuo a través de su fisonomía.

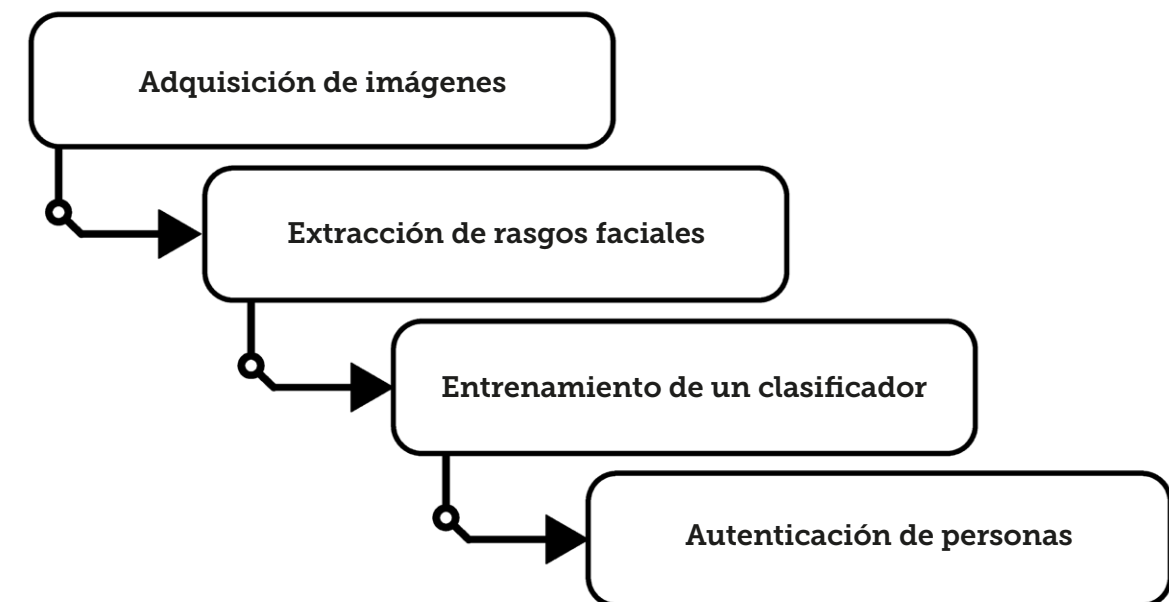
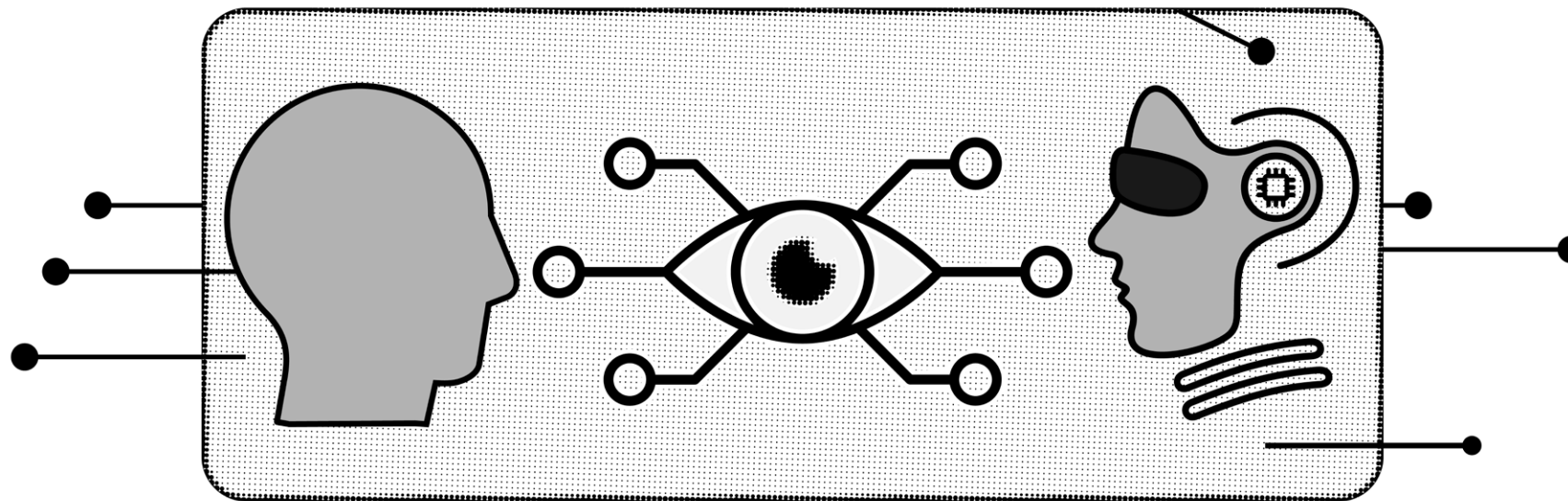


Figura 3. Diagrama general del proceso de autenticación.



AUTENTICACIÓN DE PERSONAS MEDIANTE RASGOS FACIALES

En esta área existen muchos autores con aportaciones relevantes. Uno de los pioneros es Woodrow Wilson B., quien desarrolló un sistema capaz de clasificar fotografías a partir de una tableta digital (RAND), en la que se podían ubicar en un plano las coordenadas de los ojos, la nariz, la boca y la línea del cabello. Este sistema guardaba esos puntos para usarlos después como referencia y así obtener un sistema manual que reconociera a las personas (Gates, 2011).

Con el transcurso de los años, el método se fue automatizando y se ha implementado en sistemas de seguridad que sirven para buscar delincuentes; Afra y Alhaji (2020), Awais *et al.* (2019) y Jahan *et al.* (2020) han implementado diferentes técnicas de seguridad ba-

sadas en el reconocimiento facial. Cabe mencionar que este tipo de procedimientos trabajan bajo restricciones, como que quienes son fotografiados se deben encontrar a cierta distancia de la cámara y además viendo de frente a ella. De igual manera, se deben presentar otros factores, como una buena iluminación y que la cámara tenga buena resolución, condiciones que favorecerán el proceso de autenticación.

Una de las redes sociales más usadas a nivel mundial, Facebook, ha implementado un sistema de reconocimiento de caras usando como base de entrenamiento las fotografías de perfil de los usuarios y aquéllas en las que se han etiquetado esos usuarios de manera manual. Al igual que esta plataforma,

existen teléfonos celulares cuyos sensores proyectan más de 30,000 puntos invisibles, esto con el fin de generar una imagen de profundidad (3D), una infrarroja (captada por una cámara especial) y, posteriormente, un procesador convierte estos puntos en una representación matemática que se utiliza como característica a comparar con una nueva imagen de entrada, lo que desbloquea el teléfono (Apple, 2018).

Hay ejemplos más simples donde se realiza la detección de rostros, como con la cámara de algún dispositivo que cuente con esta opción: si la persona se encuentra de frente y con una buena iluminación, se detecta dónde está el rostro, pero si se encuentra de perfil o con alguna obstrucción, el dispositivo no la puede detectar.

NICHO DE OPORTUNIDAD

Aún existen algunos problemas por resolver. Como se mencionó, existen sistemas que pueden autenticar a las personas, pero tienen una buena precisión debido a las condiciones bajo las que trabajan, y gracias a la base de imágenes que se toma como entrenamiento, ya que, en su mayoría, las bases de datos que se encuentran en la red son de gente que está a la misma distancia con respecto a la cámara (de buena calidad), totalmente de frente a ella y con una buena iluminación.

En términos generales, las etapas más importantes para una buena autenticación son la extracción de fisionomías y la implementación del clasificador. Con la primera se hace referencia a obtener información relevante de ciertos aspectos de

la cara: la forma y el tamaño, color y textura de la piel, tamaño de la nariz, ojos y boca. El segundo se encarga de generar un modelo que relaciona la información de entrada (vector característico) respecto a la de salida.

Enfocándose en extracción de peculiaridades, se cuenta con diversas técnicas que implementan algoritmos digitales para llevarla a cabo; éstos se pueden dividir en tres secciones: basados en el análisis de textura, de imágenes en 3D y en subespacios.

La primera sección, como su nombre lo dice, analiza la textura (fina, suave, áspera, gruesa), es decir, el conjunto de patrones que se presentan en una región dentro del cuadro. El análisis de esta variable se puede realizar de manera global (en toda la re-

presentación) o de manera local (en pequeñas regiones; figura 4). La textura no solamente se puede analizar en imágenes planas (2D), también se puede analizar en 3D.

El análisis de imágenes en 3D aporta otro tipo de información en comparación con el 2D. Hay individuos que tendrán diferentes formas de nariz o los labios más prominentes que otros, incluso, hay quien tiene los ojos más salidos o hundidos que los demás y estas características se pueden medir a través de la profundidad, estos detalles se pueden obtener con el 3D. De igual modo, puede servir para analizar el tamaño, la forma y la simetría de algunas partes del cuerpo.

Por último, en la tercera sección, hay algunas técnicas que hacen un cambio de

color a los retratos con el fin de obtener otro tipo de datos, como analizar los valores de los píxeles con relación a la luminosidad o a la saturación que se presenta. Por otro lado, hay otras metodologías que cambian el espacio a analizar, es decir, reducen un

conjunto de componentes (elementos dentro de la fotografía) con la finalidad de eliminar factores redundantes y así agilizar el análisis.

Tomando en cuenta las diferentes tecnologías, las particularidades que se analizan

y el tipo de valores que proporciona cada una, los sistemas de autenticación por desarrollo deben generar una respuesta en un lapso corto, con el fin de simular el proceso que realiza el cerebro al momento de autenticar a alguien conocido.

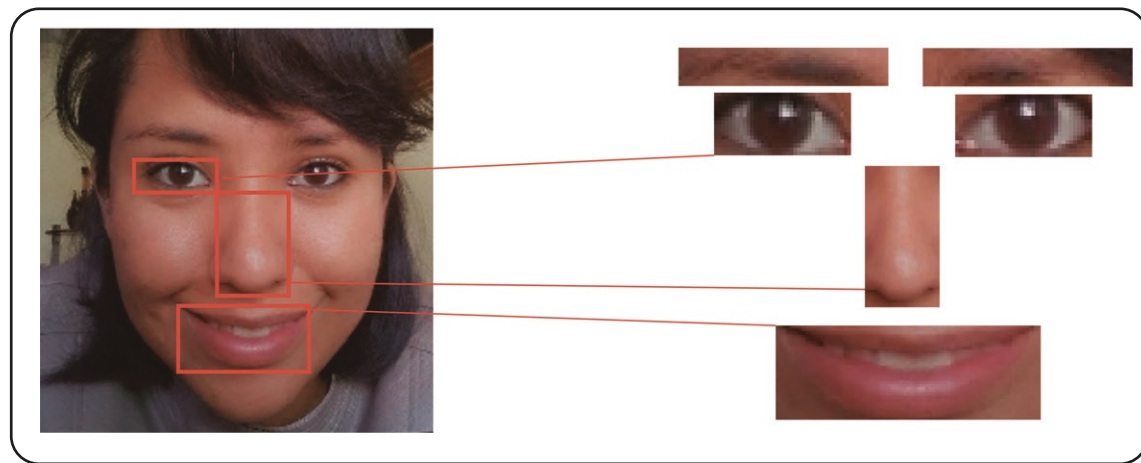


Figura 4. Análisis por regiones.

HUMANOIDES Y HUMANOS

Para que el humanoide realice la tarea de autenticación, se debe diseñar un sistema basado en los pasos mencionados, capaz de ubicar dónde se encuentran los rostros, extraer las características o los rasgos faciales más importantes, almacenar las que servirán para realizar una clasificación y posteriormente concluir con la autenticación.

Aquí es fundamental contar con un etiquetado previo, es decir, a cada uno de los cuadros adquiridos con anterioridad se le debe relacionar con una etiqueta, que en este caso sería el nombre de la o las personas a autenticar.

Hoy en día se sigue investigando cómo es posible que un RH pueda lograrlo sin la

necesidad de que los individuos se encuentren a cierta distancia, o que no necesariamente estén viendo de frente a la cámara. Aquí se toma en cuenta una serie de aspectos, como que si se está de perfil se tendría que hacer una rotación de la cara utilizando diferentes funciones, o partiendo de que ésta es simétrica, se podría trabajar con

la mitad del rostro e incluso hacer una reconstrucción en 3D. Otro punto a considerar es el diseño de un algoritmo capaz de detectar las caras en diferentes posiciones y a diferentes distancias.

Como ya se ha mencionado, lograr que un RH interactúe de forma natural con un humano plantea un sinfín de retos, en los que diferentes áreas como la visión computacional, la robótica, el procesamiento del lenguaje, entre otras, colaboran estrechamente. Respecto a las aplicaciones actuales, enfocándonos en humanoides, la empresa Hanson Robotics ha desarrollado diferentes prototipos, pero el más conocido es Sophia, que cuenta con inteligencia artificial (IA), procesamiento de datos visuales, reconocimiento y expresiones faciales, entre otras cosas. Para la etapa de autenticación utiliza una red neuronal convolucional (algoritmo basado en los impulsos neuronales biológicos). Sophia fue previamente entrenada con un conjunto de figuras etiquetadas, y con ayuda de las cámaras en su pecho y ojos puede lograr la tarea de autenticación y reconocimiento de las expresiones faciales (Hanson Robotics, 2019).

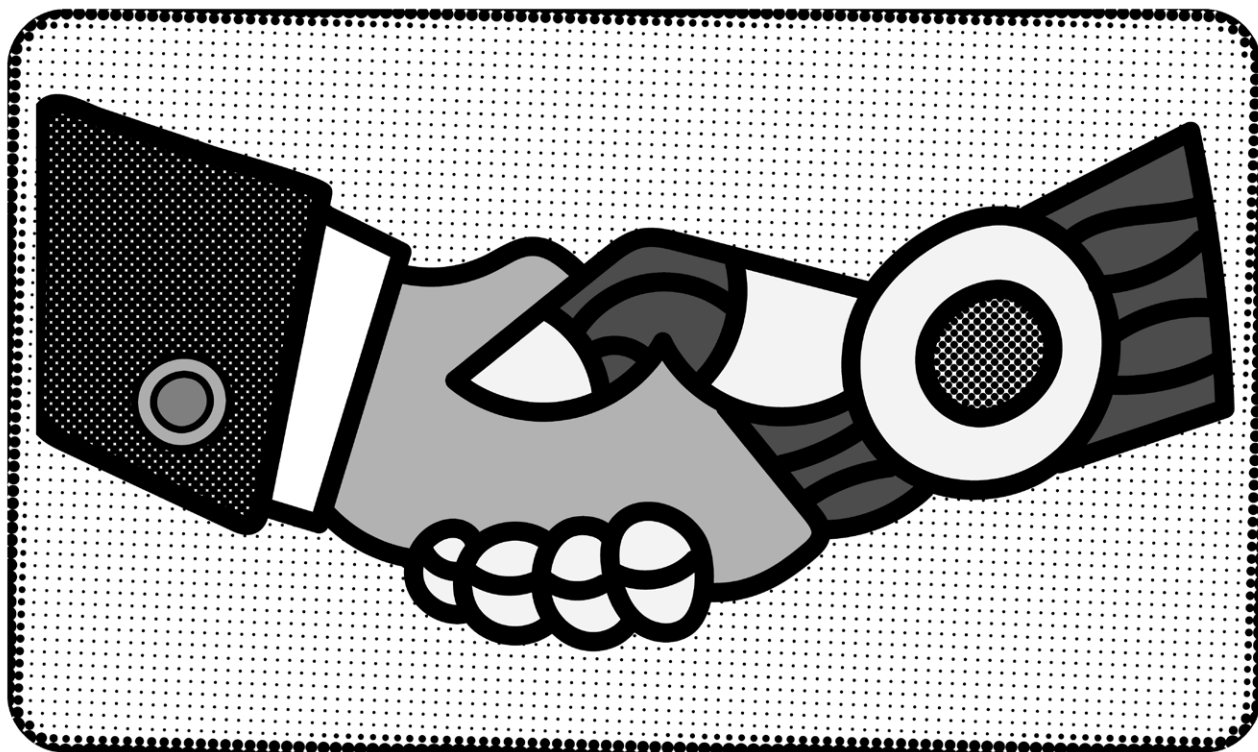
Al igual que ésta, hay empresas internacionales que han desarrollado diferentes humanoides. Como el presentado por Honda en 2000, actualizado en 2011 con sensores visuales y auditivos que le permiten reconocer los rostros y la voz (Honda, 2021). Otra muestra es DaL-e, desarrollado por Hyundai, cuyo objetivo es utilizarlo como herramienta en el servicio al cliente automatizado. DaL-e cuenta con tecnología de inteligencia artificial para el reconocimiento facial (Hyundai, 2021). Así como éstos, hay

muchos más que son comerciales o conocidos, pero hay otros que se encuentran en laboratorios de universidades o centros de investigación.

En particular, en este documento se ha presentado un panorama general de los desafíos que existen en el área de visión computacional para lograr que los RH puedan identificar y autenticar personas en su entorno cercano, así como lo hace el ser humano a través de la visión para poder interactuar de manera eficiente (figura 5).



Figura 5. Conociéndonos a través de la interacción.



CONCLUSIONES

Como se ha mencionado a lo largo de este texto, aún se sigue investigando en las diferentes áreas cómo lograr el objetivo de que un humanoide se pueda comportar lo más parecido a un humano, y es aquí donde aparece la inteligencia artificial. Actualmente, algunos sistemas de reconocimiento facial trabajan con la implementación de redes neuronales que, en conjunto con algunos algo-

ritmos, realizan el reconocimiento.

Cabe destacar que una de las metas de estos sistemas es lograr un porcentaje alto de reconocimiento, pero también hay que tener en cuenta que el tiempo de ejecución debe ser reducido, y para llevarlo a cabo es necesario diseñar un método que sea capaz de reconocer o autenticar a la gente en un lapso corto.

Con el paso de los años, la robótica seguirá desarrollando máquinas que puedan realizar un mayor número de tareas similares a las que hacen los humanos de manera cotidiana, esto no quiere decir que en el futuro puedan sustituirnos por completo, pero habrá trabajos o acciones en las que ellos obtendrán un resultado más preciso en comparación con lo realizado por nosotros.

REFERENCIAS

Afra, S., y Alhajj, R. (2020). Early warning system: From face recognition by surveillance cameras to social media analysis to detecting suspicious people. *Physica A: Statistical Mechanics and its Applications*. 540:123151.

Apple. (2018). *Acerca de la tecnología avanzada de Face ID*. Disponible en: <https://support.apple.com/es-mx/HT208108>

Awais, M., Iqbal, M.J., Ahmad, I., et al. (2019). Real-time surveillance through face recognition using HOG and feedforward neural networks. *IEEE Access*. 7:121236-121244.

Gates, K.A. (2011). *Our biometric future: Facial recognition technology and the culture of surveillance*. NYU Press: EE UU.

Hanson Robotics. (2019). *The Making of Sophia: Facial Recognition, Expressions and The Loving AI Project*. Disponible en: <https://www.hansonrobotics.com/the-making-of-sophia-facial-recognition-expressions-and-the-loving-ai-project/>

Honda. (2021). *ASIMO, el robot humanoide más avanzado del mundo*. Disponible en: <https://www.honda.mx/asimo>

Hyundai. (2021). *Hyundai Motor Group Introduces Advanced Humanoid Robot 'DAL-e'*. Disponible en: <https://www1.hyundai.news/uk/brand/hyundai-motor-group-introduces-advanced-humanoid-robot-dal-e/>

Jahan, N., Bhuiyan, P. K., Moon, P.A., et al. (2020). Real Time Face Recognition System with Deep Residual Network and KNN. In *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, IEEE, 1122-1126.

Jain, A.K., Ross, A.A., y Nandakumar, K. (2011). *Introduction to biometrics*. Springer Science & Business Media.

Wood, L.J., Zaraki, A., Robins, B., y Dautenhahn, K. (2021). Developing kaspar: a humanoid robot for children with autism. *International Journal of Social Robotics*. 13(3):491-508.

