

## Implementación del juego pares o nones en un robot social

Borrero, J.<sup>a</sup>, Arrojo, G.<sup>a</sup>, García, J.<sup>a,\*</sup>, Castillo, J.C.<sup>a</sup>, Castro-González, A.<sup>a</sup>, Salichs, M.A.<sup>a</sup>

<sup>a</sup>Departamento de Ingeniería de Sistemas y Automática, Universidad Carlos III de Madrid. Calle de Butarque, 15. Leganés, 29811. Madrid, España.

**To cite this article:** Borrero, J., Arrojo, G., García, J., Castillo, J.C., Castro, A., Salichs, M.A. 2023. Implementation of the even or odd game in a social robot.

XLIV Jornadas de Automática, 539-544. <https://doi.org/10.17979/spudc.9788497498609.539>

### Resumen

En el presente trabajo se ha desarrollado e implementado el juego *pares o nones* como una aplicación de entretenimiento en un robot social. Se ha escogido este juego al ser apto para todos los públicos, principalmente niños y personas mayores, por su simplicidad de uso y facilidad de aprendizaje. El objetivo del juego es servir tanto como aplicación de entretenimiento como para estimulación cognitiva al requerir realizar pequeñas operaciones matemáticas. Para llevar a cabo esta aplicación, por un lado, se ha utilizado modelos de inteligencia artificial para la detección de las manos de los usuarios y se han realizado las operaciones de procesado correspondientes para obtener la información necesaria en el juego. Por otro lado, se ha implementado la habilidad en el robot que le permite al usuario interactuar con el mismo. Tanto el módulo de detección como la habilidad se han integrado en la arquitectura completa del robot.

**Palabras clave:** Robótica, Percepción y detección, Estimulación cognitiva, Interacción humano-robot, Entretenimiento

### Implementation of the even or odd game in a social robot

#### Abstract

In this work, the game *even or odd* has been developed and implemented as an entertainment application in a social robot. This game has been chosen to be suitable for all audiences, mainly children and elderly people, for its simplicity of use and ease of learning. It serves two purposes: entertainment and cognitive stimulation by performing small mathematical operations throughout its course. To carry out this application, on the one hand, artificial intelligence models have been used to detect the users' hands, and the corresponding processing operations have been carried out to obtain the necessary information in the game. On the other hand, the ability has been implemented in the robot that allows the user to interact with it. The detection module and the skill have been integrated into the robot architecture.

**Keywords:** Robotics, Perception and sensing, Cognitive stimulation, Human-robot interaction, Entertainment

### 1. Introducción

La robótica social está experimentando un auge en nuestra sociedad debido al aumento del rango de las aplicaciones disponibles (Núñez et al., 2011). Se busca que la interacción con este tipo de robots se lleve a cabo con naturalidad y fluidez. Se utilizan en diversas aplicaciones como la educación, la salud, el entretenimiento, la comunicación y las tareas que requieren de trabajo en equipo (Breazeal et al., 2016).

En este contexto, se incluyen los juegos de entretenimien-

to, como puede ser el *pares o nones*. Estos juegos ofrecen una forma amena y accesible de interactuar con los robots sociales. Al implementar juegos simples y populares en estos robots, los usuarios pueden iniciar la interacción sencillamente.

La implementación de juegos tradicionales en la robótica social presenta algunas particularidades. Por un lado, son juegos fáciles de entender y, en caso de no conocer sus reglas, de aprender a jugarlos (Soute and Markopoulos, 2007). Esto reduce las posibles dificultades para los nuevos usuarios, lo que puede resultar de interés para aquellos que no están familiari-

\*Autor para correspondencia: [jesusgar@ing.uc3m.es](mailto:jesusgar@ing.uc3m.es)  
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

zados con la tecnología. Por otro lado, los juegos permiten que tanto el usuario como el robot participen en la interacción activamente, lo que permite al usuario percibir la vivacidad del robot (Xin and Sharlin, 2007).

Para que el robot pueda participar debidamente en el juego debe ser capaz de obtener datos de su entorno y del usuario, utilizando para ello sistemas de percepción que sean capaces de detectar a los usuarios y extraer la información pertinente a través de la cámara. Los modelos de inteligencia artificial permiten detectar las zonas de interés en la imagen, como pueda ser la cara de la persona con la que el robot interactúa (Rowley et al., 1998), o las manos de la misma (Zhang et al., 2020). Con algoritmos de procesamiento de imágenes, se pueden analizar las regiones que contienen las manos y realizar el recuento de dedos para determinar si la suma es par o impar. Por ejemplo, en el trabajo de (Gualdrón et al., 2013), se hace uso de la visión artificial y el reconocimiento de patrones para poder detectar y diferenciar rostros de diferentes usuarios para ser usado en ámbitos de seguridad.

Un desarrollo reciente integra el juego del *Tangram* en un robot social (Menendez et al., 2021). Los autores que implementaron el juego del Tangram proponen el uso de un sistema de percepción capaz de detectar el tablero y las piezas del juego en tiempo real para guiar al usuario mientras realizaba las figuras. Del mismo modo, en Tena Lara (2019) se implementa el juego del *ahorcado* como aplicación de entretenimiento para un robot social. Un último ejemplo es la implementación de un juego de cartas donde el robot juega competitivamente a la vez que evalúa la situación y responde emocionalmente de manera natural (Correia et al., 2017).

Este trabajo se divide en las siguientes partes: En la sección 2 se describen las herramientas utilizadas, incluyendo la plataforma robótica y la arquitectura software empleada para el desarrollo del detector y la aplicación. En la sección 3 se explican los diferentes algoritmos y modelos utilizados para la detección de las manos del usuario y el procesamiento de imágenes realizado. A continuación, en la sección 4 se detalla el funcionamiento de la habilidad integrada en el robot. Finalmente, en la sección 5 se recogen las principales conclusiones a partir del trabajo realizado.

## 2. Herramientas utilizadas

La plataforma sobre la que se ha realizado el desarrollo e integración del detector de manos y la habilidad es el robot social Mini (Salichs et al., 2020). Este robot está diseñado para realizar tareas de estimulación cognitiva con niños y personas mayores. Para ello, está dotado de aplicaciones y juegos con los que el usuario puede interactuar, ofreciendo facilidades y entretenimiento. Algunas de estas aplicaciones incluyen la reproducción de imágenes y vídeos, relatar noticias de actualidad o aplicaciones de entretenimiento como el juego del bingo, entre otras (Velázquez-Navarro et al., 2019). Su apariencia física se caracteriza por tener un aspecto amigable (ver figura 1). Además, el color del revestimiento de peluche del robot se puede personalizar a gusto del usuario para diferenciarlo del resto.



Figura 1: Robot social Mini.

En lo referido a su diseño a nivel de *hardware*, el robot Mini cuenta con un cuerpo blando, articulaciones en brazos, cabeza y torso, ojos expresivos para mostrar sentimientos y gestos y luces de colores en el corazón, mejillas y boca. También dispone de una tableta auxiliar para mostrar contenido multimedia, altavoces, micrófono interno, sensores táctiles y cámaras incorporadas. Estos elementos dotan al robot de la capacidad de interacción multimodal para establecer una comunicación natural con el usuario.

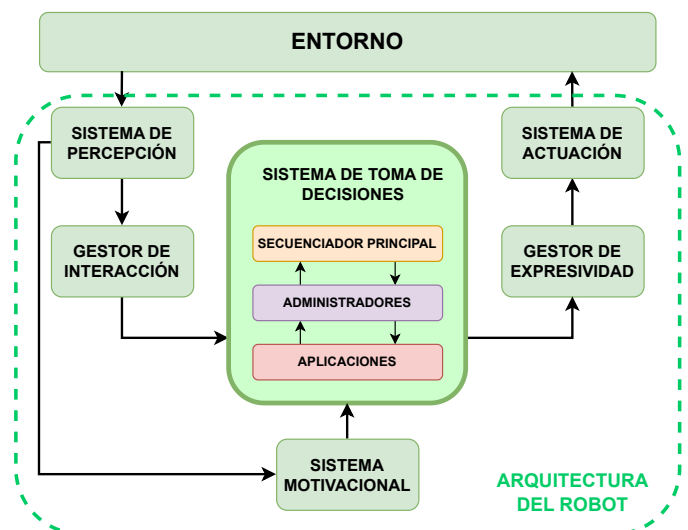


Figura 2: Arquitectura de Mini.

Por otro lado, la arquitectura software de Mini utiliza como marco de trabajo *Robot Operating System (ROS)* (Quigley et al., 2009). Como se puede observar en la figura 2, la arquitectura está compuesta de varios módulos interconectados: Por

un lado se encuentra el sistema de percepción (Alonso Martín et al., 2017), encargado de adquirir y preprocesar la información del entorno y del usuario, a nivel visual y táctil. Dicha información es enviada a los módulos de gestión de la interacción (Rodicio, 2021) y Sistema de Toma de Decisiones (DMS) (Maroto-Gómez et al., 2018), encargados de gestionar la información previa y generar la interacción con el usuario y la toma de decisiones del robot. Finalmente, se encuentran el gestor de la expresividad (Maroto-Gómez et al., 2022) y el sistema de actuación encargados de gestionar los gestos y emociones que el robot expresa, así como su movimiento.

En lo referente a las librerías principales utilizadas en el trabajo desarrollado, para el módulo de detección de manos (ver sección 3 se han utilizado los modelos de visión de *Mediapipe* (Lugaresi et al., 2019) y para el procesamiento de imágenes la librería *OpenCV* (Bradski and Kaehler, 2008). *Mediapipe* es un *framework* de bajo nivel compuesto por modelos preentrenados de visión (detección y clasificación), lenguaje natural y audio basados en *machine learning* y *deep learning*. La optimización en el consumo de recursos permite su ejecución en tiempo de real durante la interacción con el usuario. La biblioteca de procesamiento de imágenes, *OpenCV*, está formada por funciones y algoritmos que permiten manipular, analizar y filtrar la información de interés dentro de una imagen. En el trabajo realizado se ha utilizado para procesar las imágenes provenientes de la cámara y adaptarlas al modelo de detección de manos. También se ha utilizado para manipular la información del detector y mostrar la información por pantalla.

### 3. Detección y clasificación de las manos del usuario

El módulo de detección de manos se encarga de proporcionarle a la habilidad del robot el número de dedos extendidos por el usuario durante la interacción. Para reducir el consumo de CPU, dicho módulo permanece inactivo hasta que la habilidad del robot le envía la señal de activación una vez llega el momento de mostrar las manos en el juego. El algoritmo de *Mediapipe* utilizado está compuesto de dos modelos de detección: detector de la palma de la mano y detector de los puntos de referencia característicos de la mano. La detección funciona en dos fases para optimizar recursos, y solo localizar los puntos de referencia cuando se está seguro de haber detectado alguna de las manos en la imagen. El modelo utilizado está preentrenado con 30.000 imágenes capturadas del mundo real e imágenes de manos sintéticas renderizadas sobre fondos variados.

En la figura 3 se puede observar las operaciones que se realizan en la imagen para calcular el número de dedos extendidos. En primer lugar, el modelo de detección de la palma de la mano ubica en la **imagen RGB** las manos del usuario a través del recuadro delimitador en la **Región de Interés (ROI)** que contiene cada mano por separado. Además, dicho modelo diferencia si se está mostrando la mano izquierda, derecha o ambas.

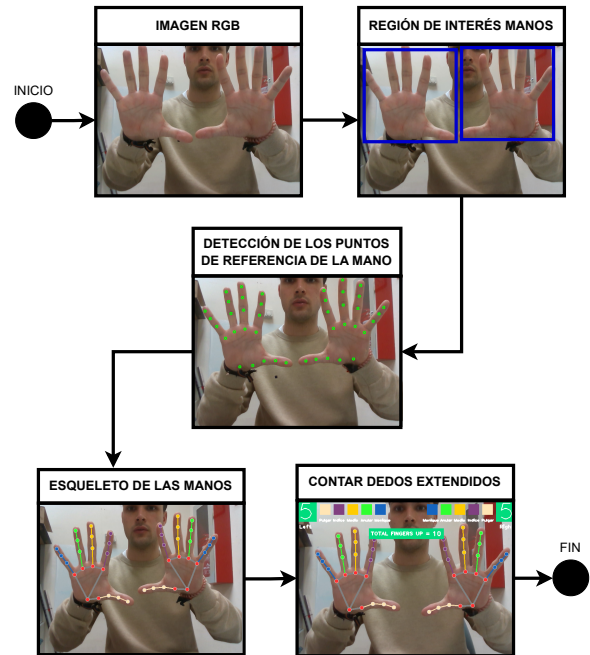


Figura 3: Detección de manos y conteo de dedos.

Una vez se detecta en la imagen la presencia de las manos del usuario se realiza la inferencia del segundo modelo sobre la ROI de cada mano recortada obteniendo los **21 puntos de referencia de la mano** (ver figura 4). A continuación, se realiza el seguimiento de los puntos obtenidos para calcular la nueva ubicación de la mano en cada fotograma. En caso de que no sea posible calcular los puntos característicos en el fotograma consecutivo utilizando la ROI del fotograma anterior, se reactiva el detector de la palma de la mano para localizar la nueva posición de las manos. Esto se hace para optimizar recursos dado que la detección de las manos es computacionalmente más costosa.

Cada punto detectado tiene un valor de confianza asociado. Se aplica un umbral para considerarlo como válido (0.5 por defecto). Por lo tanto, siempre y cuando los 21 puntos de referencia estén por encima del umbral establecido, se dibuja el **esqueleto de la mano**. En caso contrario, uno o varios puntos queden por debajo del umbral, las manos se considerarán no detectadas y se reactivará el detector de la palma de la mano.

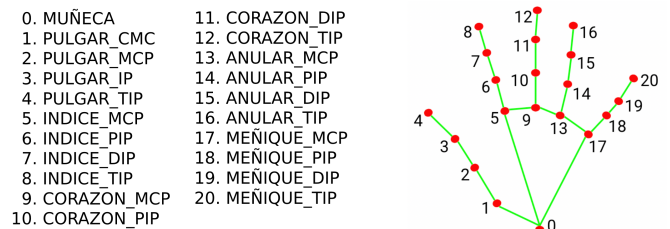


Figura 4: Puntos de referencia de la mano: <sup>1</sup>

Por último, es necesario **contar los dedos extendidos** para su posterior uso. Para determinar si un dedo está extendido o contraído, en el caso del pulgar se emplea el teorema del co-

<sup>1</sup>[https://developers.google.com/mediapipe/solutions/vision/hand\\_landmarker](https://developers.google.com/mediapipe/solutions/vision/hand_landmarker)

seno (ecuación 1), donde ( $\alpha$ ) es el ángulo que determina cómo de abierto está el pulgar (ver figura 5). Se establece como valor umbral para el ángulo 150 grados, valor a partir del cual se considera que el dedo está extendido. Además,  $l_1$ ,  $l_2$  y  $l_3$  son las distancias entre los puntos 2-3, 2-4 y 3-4 respectivamente, y se corresponden con las partes del pulgar. Estas medidas se calculan a partir de la distancia euclídea (ecuación 2), donde  $x$ , y corresponden a las coordenadas (en píxeles) de los puntos en la imagen RGB.

$$\alpha = \arccos\left(\frac{l_1^2 + l_3^2 - l_2^2}{2l_1l_3}\right) \quad (1)$$

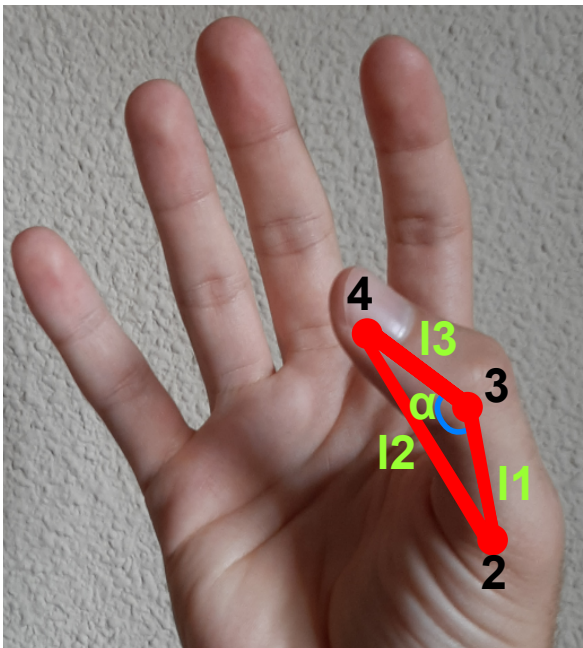


Figura 5: Detección del pulgar extendido.

$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (2)$$

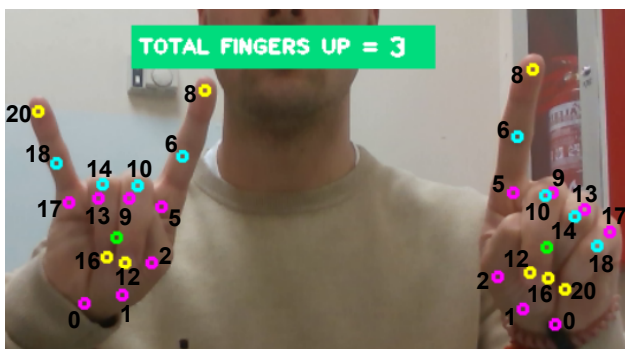


Figura 6: Puntos de referencia de las manos. Verde: centroide de la palma de la mano. Morado: 0, 1, 2, 5, 9, 13, 17. Azul: 6, 10, 14, 18. Amarillo: 8, 12, 16, 20.

Para el resto de dedos, atendiendo a la figura 6: primero se calcula el centroide de la palma de la mano (*punto verde*) utilizando todos los puntos que componen la palma de la mano

(*puntos rosas*). El centroide se utiliza como referencia para el cálculo de distancias. Para determinar si los dedos están expandidos o contraídos, se comparan dos medidas: si la distancia entre el *punto amarillo* de cada dedo y el centroide es menor que la distancia entre el *punto azul* de ese mismo dedo y el centroide significa que el dedo está contraído. Para calcular la distancia se utiliza la distancia euclídea (ecuación 2). En caso de que uno o varios dedos no se detecten se considerarán contraídos.

Finalmente, se suma el número de dedos extendidos en ambas manos y se le envía esa información a la habilidad del robot en el momento en el que el módulo de detección se encuentra activo.

#### 4. Desarrollo de la habilidad del robot

El juego del pares o nones es un juego de azar en el que dos jugadores eligen si quieren ser pares/nones antes de iniciar la partida, para posteriormente realizar una cuenta regresiva y mostrar simultáneamente las manos previamente escondidas con los dedos extendidos. A continuación, se suman todos los dedos extendidos y, si la suma de ambos es par, gana el jugador que eligió pares, en caso contrario gana el jugador que es nones.

A la hora de diseñar la habilidad del robot se ha replicado la estructura del juego real, añadiendo una etapa inicial en la que el robot le preguntará al usuario si conoce el juego. En caso de que nunca haya jugado le explicará las instrucciones. Como los usuarios finales está planteado que sean niños o personas mayores, el robot en todo momento guiará al usuario a lo largo de la interacción indicándole lo que debe hacer tanto por comunicación verbal como visual a través de la tableta. En el diseño de la aplicación la cuenta regresiva para mostrar las manos siempre la realizará el robot dado que es el que está guiando al usuario durante el transcurso de la aplicación.

En la figura 7 se puede observar las diferentes etapas por las que pasa la habilidad del robot durante el transcurso del juego: al inicio del juego, el robot le pregunta al usuario si conoce el juego y sus reglas o por el contrario desea que le recuerde las instrucciones. Antes de empezar a jugar, toca escoger quién es pares y quién es nones y, para ello, existen dos posibilidades que suceden al azar: el robot toma la iniciativa de escoger ser pares o nones de manera aleatoria, o el usuario es el que elige la opción que desee. Se juega al mejor de tres rondas de manera que no se pueda empatar. Una vez se haya establecido quién es pares y quién es nones se pasará a la primera ronda.

En cada ronda el robot realizará una cuenta regresiva y cuando de la señal: “¡Ya!”, tanto el usuario como el robot mostrarán las manos a la vez. Existe la posibilidad de mostrar números con una sola mano o con las dos a la vez, por lo que el rango de valores para cada usuario está entre el 1 y el 10, ambos incluidos. Dado que Mini no dispone de dedos físicos, además de comunicarlo verbalmente, mostrará en una tableta auxiliar el número de dedos que ha escogido sacar como se observa en la figura 8. En el momento en el que el usuario tiene que mostrar las manos, es cuando se activa el detector de manos explicado en la sección 3. El número de dedos mostrados de ambos jugadores se sumará y si el número es par ganará el que juega pares y si el número es impar ganará el que juega nones.

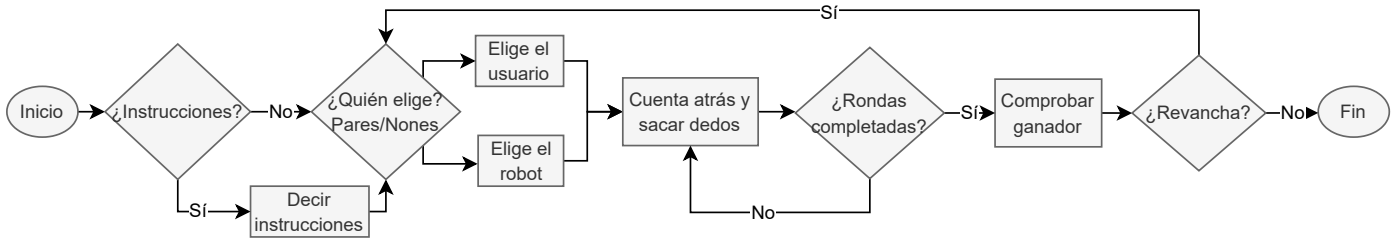


Figura 7: Diagrama de flujo de la aplicación.



Figura 8: Dedos mostrados por el robot en la tableta auxiliar.

Hasta que el usuario no muestre las manos al robot en el momento indicado, el robot no mostrará su elección. En caso de que el usuario se olvide de mostrar las manos, o que el robot no sea capaz de detectarlas (cero dedos detectados), le avisará por voz y volverá a realizar la cuenta atrás repitiendo la ronda. Este proceso se repite hasta que haya un jugador victorioso. Tras finalizar el juego, el robot preguntará al usuario si quiere volver a jugar y optar por la revancha.

El siguiente vídeo muestra un ejemplo de interacción real en el que un usuario juega una partida completa con el robot: <https://youtu.be/2XCq4j4WfuI>.

## 5. Conclusiones

En este trabajo se ha desarrollado e integrado un juego de entretenimiento cotidiano en un robot social. Para ello, se ha implementado por un lado un módulo de percepción encargado de detectar las manos del usuario y, por otro lado, la habilidad necesaria para jugar con los usuarios.

El módulo de detección utilizado, compuesto por el detector de la palma y el detector de los puntos de referencia, es capaz de identificar las manos del usuario en la imagen, diferenciar entre ambas manos y detectar las posiciones de cada uno de los dedos para un único usuario situado frente al robot.

La habilidad del robot permite que la interacción se realice de manera natural y se ha optimizado el consumo de recursos en robot activando el detector únicamente cuando la habilidad lo indica, a la hora de mostrar las manos.

La aplicación además de ser un juego de entretenimiento tiene fines de estimulación cognitiva y física al requerir realizar pequeñas operaciones aritméticas al mismo tiempo que el usuario ejercita la motricidad fina para mostrar los dedos.

## Agradecimientos

La investigación que ha conducido a estos resultados ha recibido financiación de los proyectos: «RoboCity2030-DIHC-M», Madrid Robotics Digital Innovation Hub, S2018/NMT-4331, financiado por «Programas de Actividades I+D en la Comunidad de Madrid» y cofinanciado por el Fondo Social Europeo (FSE) de la EU; «Robots sociales para mitigar la soledad y el aislamiento en mayores» (SoRoLi), PID2021-123941OA-I00, financiado por la Agencia Estatal de Investigación (AEI), Ministerio Español de Ciencia e Innovación; «Robots sociales para reducir la brecha digital de las personas mayores» (SoRo-Gap), TED2021-132079B-I00, financiado por la Agencia Estatal de Investigación (AEI); «Mejora del nivel de madurez tecnológica del robot Mini» (MeNiR), PDC2022-133518-I00, financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea NextGenerationEU/PRTR.

## Referencias

- Alonso Martín, F., Castillo, J.C., Castro González, Á., Gamboa, J.J., Maroto, M., Marqués Villarroya, S., Pérez Vidal, A.J., Salichs, M.Á., et al., 2017. Percibiendo el entorno en los robots sociales del roboticslab. *Actas de las XXXVIII Jornadas de Automática*.
- Bradski, G., Kaehler, A., 2008. *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media, Inc.ª.
- Breazeal, C., Dautenhahn, K., Kanda, T., 2016. *Social robotics*. Springer handbook of robotics, 1935–1972.
- Correia, F., Alves-Oliveira, P., Ribeiro, T., Melo, F., Paiva, A., 2017. A social robot as a card game player, in: *Proceedings of the AAAI conference on artificial intelligence and interactive digital entertainment*, pp. 23–29.
- Gualdrón, O.E., Suárez, O.M.D., Rojas, M.A.C., 2013. Diseño de un sistema de reconocimiento de rostros mediante la hibridación de técnicas de reconocimiento de patrones, visión artificial e ia, enfocado a la seguridad e interacción robótica social. *Mundo Fesc* 3, 16–28.
- Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.L., Yong, M.G., Lee, J., et al., 2019. Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172*.
- Maroto-Gómez, M., Castro-González, Á., Castillo, J.C., Malfaz, M., Salichs, M.A., 2018. A bio-inspired motivational decision making system for social robots based on the perception of the user. *Sensors* 18, 2691.
- Maroto-Gómez, M., Villarroya, S.M., Malfaz, M., Castro-González, Á., Castillo, J.C., Salichs, M.Á., 2022. A preference learning system for the autonomous selection and personalization of entertainment activities during human-robot interaction, in: *2022 IEEE International Conference on Development and Learning (ICDL)*, IEEE. pp. 343–348.
- Menendez, C., Marques-Villarroya, S., Castillo, J.C., Gamboa-Montero, J.J., Salichs, M.A., 2021. A computer vision-based system for a tangram game in a social robot, in: *Ambient Intelligence—Software and Applications: 11th International Symposium on Ambient Intelligence*, Springer. pp. 61–71.
- Núñez, P., Bustos, P., Jaramillo, E., Bachiller, P., García-Varea, I., 2011. Robots sociales para la mejora de la calidad de vida de las personas dependientes, in: *VI Congreso Iberoamericano de Tecnologías de Apoyo a la Discapacidad IBERDISCAP*, pp. 94–103.

- Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Wheeler, R., Ng, A.Y., et al., 2009. Ros: an open-source robot operating system, in: ICRA workshop on open source software, Kobe, Japan. p. 5.
- Rodicio, E.F., 2021. Human-robot interaction architecture for interactive and lively social robots. Ph.D. thesis. Universidad Carlos III de Madrid.
- Rowley, H.A., Baluja, S., Kanade, T., 1998. Neural network-based face detection. *IEEE Transactions on pattern analysis and machine intelligence* 20, 23–38.
- Salichs, M.A., Castro-González, Á., Salichs, E., Fernández-Rodicio, E., Maroto-Gómez, M., Gamboa-Montero, J.J., Marques-Villarroya, S., Castillo, J.C., Alonso-Martín, F., Malfaz, M., 2020. Mini: a new social robot for the elderly. *International Journal of Social Robotics* 12, 1231–1249.
- Soute, I., Markopoulos, P., 2007. Head up games: the games of the future will look more like the games of the past, in: *Human-Computer Interaction–INTERACT 2007: 11th IFIP TC 13 International Conference*, Rio de Janeiro, Brazil, September 10-14, 2007, Proceedings, Part II 11, Springer. pp. 404–407.
- Tena Lara, F., 2019. Desarrollo e implementación del juego del Ahorcado en el robot Mini. B.S. thesis.
- Velázquez-Navarro, E., González-Díaz, S., Alonso-Martina, F., Castiello, J.C., Castro-González, A., Malfaza, M., Salichs, M.A., 2019. El robot social mini como plataforma para el desarrollo de juegos de interacción multimodales, in: *Jornadas Nacionales de Robotica, Spanish Robotics Conference*.
- Xin, M., Sharlin, E., 2007. Playing games with robots-a method for evaluating human-robot interaction, in: *Human Robot Interaction*, IntechOpen.
- Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C.L., Grundmann, M., 2020. Mediapipe hands: On-device real-time hand tracking. arXiv preprint arXiv:2006.10214 .