



Interfaz Humano-Máquina: Control de Cobots Usando Señas Estáticas de la Mano

Ángel Gabriel Salinas Martínez ^{1,*}, Eduardo Ulises Conde Cruz ¹ y Ángel Iván García Moreno¹

¹ Centro de Ingeniería y Desarrollo Industrial. Playa Pie de la Cuesta 702, Desarrollo San Pablo, Querétaro, México. 76125.

* Autor de correspondencia: angel.salinas@cidesi.edu.mx

Resumen: En la actualidad, la mayoría de las interfaces humano-máquina (HMI, por sus siglas en inglés) se basan en medios físicos, como teclados y pantallas táctiles. Sin embargo, existen otras alternativas para interactuar con las máquinas que resultan más naturales e intuitivas, una de ellas es el control mediante señas de la mano. Una de las principales ventajas de esta alternativa es que no requiere conocimientos técnicos sobre el método de funcionamiento de una máquina, sino que basta con disponer de un diccionario de señas para garantizar una curva de aprendizaje rápida y sencilla. La interacción mediante señas puede realizarse sin necesidad de estar físicamente en el mismo entorno que una máquina, lo que garantiza una mayor seguridad para el usuario. Este trabajo propone una interfaz HMI utilizando el Leap Motion Controller (LMC), un dispositivo que utiliza cámaras infrarrojas para seguir los movimientos de la mano. También se propone un diccionario de señas estáticas que se basa en la clasificación y cuantificación de los dedos extendidos. La precisión alcanzada en el reconocimiento de las señas definidas en este diccionario fue del 96.75% y del 95.77% para la mano izquierda y la mano derecha respectivamente. A partir del reconocimiento de señas estáticas con LMC, se dan instrucciones a tres robots colaborativos (cobots) para que ejecuten tareas específicas en un entorno de fabricación de placas de circuito impreso (PCBs, por sus siglas en inglés).

Palabras clave: Reconocimiento de señas, Leap Motion Controller, Interfaz Humano-Máquina.

1. Introducción

Las HMI se han convertido en un componente esencial en campos como la robótica y la automatización industrial, ya que facilitan la interacción entre humanos y máquinas. Una interacción segura es la clave para mejorar la eficiencia y flexibilidad en los entornos de manufactura. En este contexto, los cobots han ganado relevancia debido a su capacidad para trabajar junto a los humanos sin representar riesgo alguno. Sin embargo, las HMI tradicionales para interactuar con un cobot dependen de pantallas táctiles y requieren conocimientos técnicos sobre su programación. En ese sentido las HMI basadas en señas representan una alternativa que además de incrementar la seguridad al permitir manipular a un robot sin siquiera tocarlo, también permite una interacción fluida y natural que no depende de conocimientos técnicos complejos.

El control mediante señas de la mano puede lograrse mediante diversas tecnologías que presentan enfoques particulares para captar el movimiento de las extremidades del cuerpo humano. Uno de los enfoques ampliamente utilizados en el estado del arte es mediante dispositivos de visión, los cuales permiten un seguimiento preciso y no invasivo de las manos del usuario. El LMC entra en esta categoría al emplear cámaras estereoscópicas y sensores de espectro infrarrojo para detectar variaciones térmicas y movimientos en condiciones de iluminación variable. Sin embargo, también existen otros enfoques que se basan en el uso de sensores inerciales compuestos por Unidades de Medición Inerciales (IMUs, por sus siglas en inglés) o sensores de electromiografía (EMG, por sus siglas en inglés) compuestos por electrodos que miden los pulsos eléctricos de los músculos. En el presente trabajo, se describe el desarrollo de una HMI basada en señas de la mano utilizando el LMC,



diseñada para controlar los movimientos de tres cobots dentro de una celda de manufactura de PCBs. El trabajo comienza con un análisis del estado del arte, donde se revisan las investigaciones recientes sobre el control de cobots mediante señas. Luego, se presenta la metodología utilizada, detallando el diseño del sistema, la integración del hardware y software, y los algoritmos empleados para el reconocimiento de señas, así como la comunicación con los cobots a través del protocolo MQTT. A continuación, se muestran los resultados, destacando la eficiencia y precisión del sistema. Finalmente, se presentan las conclusiones, discutiendo los logros obtenidos, concluyendo que la HMI basada en gestos es una herramienta eficaz para mejorar la interacción humano-cobot.

2. Trabajos relacionados

El LMC ha sido evaluado e implementado en diversas aplicaciones de desarrollo e investigación en los últimos años. Un estudio ha demostrado que puede rastrear manos más allá de su rango máximo de 60 cm, alcanzando en algunos casos hasta 100 cm [1]. Además, se ha demostrado su efectividad para el control de robots industriales y manipuladores colaborativos mediante técnicas basadas en cinemática directa e inversa, como se observa en el control de un robot Delta [2] y la mano robótica DLR-HIT II [3]. En otra investigación se diseñó una interfaz de software intuitiva que permite a un usuario controlar un robot ABB IRB-140 utilizando el LMC para el reconocimiento de señas y manteniendo un espacio de trabajo seguro tanto para el robot como para el usuario [4]. Otro trabajo implementó el reconocimiento de señas con el LMC para realizar operaciones tales como encendido y apagado de un robot Phantom X Pincher [5]. Otras aplicaciones incluyen el control de interfaces gráficas de usuario de sistemas operativos basados en Linux [6]. También se ha empleado el LMC para controlar un robot colaborativo UR5e mediante señas [7]. De manera similar, en [8] emplearon el LMC junto con el brazalete Myo para reconocer las señas de la mano en tiempo real, los cuales aplicaron para controlar las acciones de un robot KUKA youBot. En [9] propusieron un marco de interacción hombre-máquina basado en gestos en el que un robot KUKA iiwa ayuda a un operador humano entregando herramientas y piezas, y sujetando objetos en una operación de montaje. [10] propusieron un marco espacio-temporal STGauntlet para detectar tipos de señas en tiempo real junto con su trayectoria de seguimiento 3D utilizando múltiples sensores de movimiento portátiles y en [11] utilizaron el mapeo 3D de señas para manejar un brazo robótico en beneficio de personas con discapacidad.

Así mismo, se han explorado métodos avanzados como la combinación de redes generativas antagónicas (GAN, por sus siglas en inglés) con redes convolucionales (CNN, por sus siglas en inglés) para la clasificación de señas [12], logrando mejoras significativas frente a modelos tradicionales. De igual manera, se han propuesto sistemas de control por señas alternativos al uso exclusivo de visión, como el empleo de radares de ondas milimétricas [13]. En el contexto de escenarios de control más complejos, que en ocasiones excepcionales exigen el uso de múltiples dispositivos, [14] desarrollaron un sistema de reconocimiento de señas basado en múltiples LMC, utilizando redes neuronales recurrentes con memoria a largo y corto plazo (RNN-LSTM, por sus siglas en inglés) para la teleoperación de robots quirúrgicos. La técnica demostró un alto nivel de precisión de reconocimiento y velocidad de inferencia. De manera similar, en [15] combinaron datos de radar y visión, mejorando la adaptabilidad y robustez del sistema a través de una red de fusión deformable con RNN-LSTM. Este enfoque de fusión de dispositivos es particularmente ventajoso en escenarios donde las condiciones ambientales son variables, abordando de manera efectiva las restricciones asociadas con los sistemas que dependen exclusivamente de un solo dispositivo.

3. Materiales y Métodos

La metodología propuesta describe el desarrollo de una HMI que permite la interacción mediante señas estáticas con los cobots que integran a una celda de manufactura de PCBs. La metodología comprende cinco niveles (ver Figura 1) los cuales se describen a continuación:

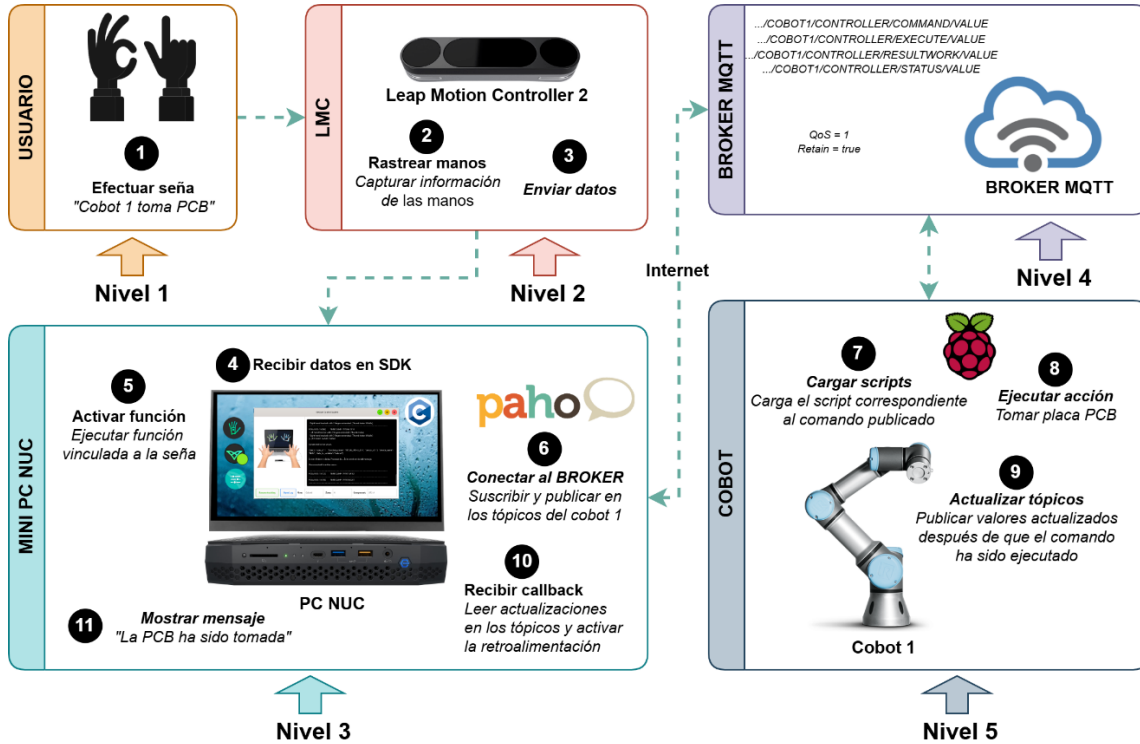


Figura 1. Metodología para el desarrollo de la HMI basada en señas.

El primer nivel corresponde al usuario que efectúa una combinación de señas con ambas manos para iniciar una interacción. En el segundo nivel se emplea un dispositivo LMC para rastrear la posición de las manos en un espacio tridimensional. En el tercer nivel se emplea una mini PC NUC con las dependencias de software detalladas en la Tabla 1. Entre estas dependencias se encuentra el Kit de Desarrollo de Software (SDK, por sus siglas en inglés), el cual ya contiene algoritmos que procesan los datos de posición para identificar y clasificar los dedos extendidos. Adicionalmente, se añadió una funcionalidad que permite cuantificar el número de dedos activos en cada interacción, contemplando ambas manos. En algunos casos también se mide la distancia entre las puntas de los dedos para detectar señas más complejas. Esta información se utiliza para definir las señas en el código fuente, de modo que, al ser reconocidas, se envía un mensaje al siguiente nivel a través del protocolo de Transporte de Telemetría de Colas de Mensajes (MQTT, por sus siglas en inglés). En el cuarto nivel un servidor, conocido como *broker* MQTT, permite la comunicación bidireccional entre los cobots y la mini PC NUC. El *broker* gestiona un árbol de tópicos que coordina y controla las acciones de los cobots en función de los mensajes que llegan desde la mini PC NUC. Finalmente, en el quinto nivel se encuentran los cobots suscritos a los tópicos del *broker* MQTT, lo que les permite recibir datos operativos que activan acciones específicas. De esta manera, cada vez que el usuario efectúa una combinación de señas en el primer nivel, los cobots desencadenan la acción correspondiente de forma inmediata.

Tabla 1. Dependencias de software para el correcto funcionamiento de la HMI.

Dependencias de software	Versión
Microsoft Visual C++ 2013 Redistributable Packages	12.0.3
Ultra Leap Gemini Development Package	5.13.2

Eclipse Paho MQTT C	1.3.12
OpenSSL library	1.1.1
CMake tools	3.27.3
GTK	4

4. Resultados

A continuación, se presentan y discuten los resultados obtenidos en el desarrollo de la HMI basada en señas. En primer lugar, se describe la estrategia para la definición de un diccionario de señas y posteriormente se discute la efectividad del algoritmo para el reconocimiento de las mismas.

Se definió un diccionario de señas (ver Figura 2) basado en una estrategia que asigna una seña específica a cada una de las estaciones que forman parte del proceso de fabricación de una PCB. Del mismo modo se asignan señas para representar las acciones que los cobots pueden realizar en ellas. Las estaciones se identifican con la mano derecha y están numeradas de acuerdo al orden en el que participan, por ejemplo, la primera estación se identifica con la seña “Uno”, la segunda estación con “Dos” y así sucesivamente. Las acciones se indican con la mano izquierda y permiten a los cobots moverse entre las estaciones y tomar o colocar una PCB.

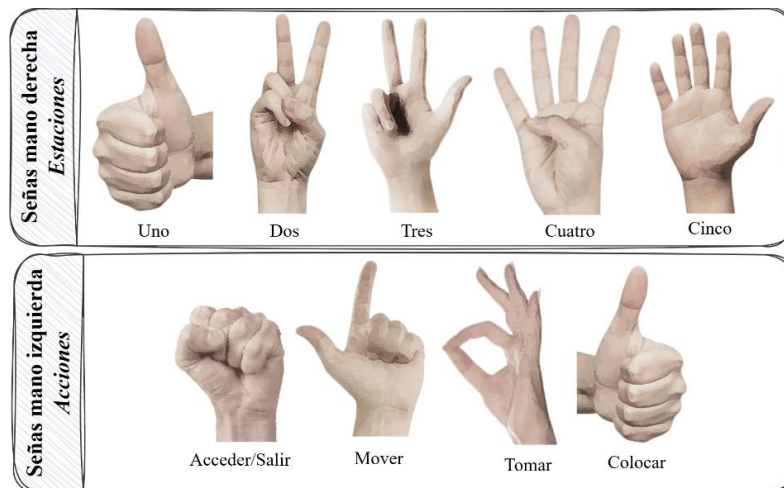


Figura 2. Diccionario de señas.

Para optimizar el diccionario de señas se propuso dividir la celda de manufactura en tres zonas (ver Figura 3), permitiendo que los tres cobots utilicen las mismas señas. La seña "Acceder" combinada con las señas "Uno", "Dos" o "Tres" indica el acceso a cada zona respectivamente. Una vez que se accede a alguna zona se puede manipular el cobot efectuando una combinación de señas con ambas manos, por ejemplo, “Tomar” la PCB de la estación “Uno”, “Mover” el cobot a la estación “Dos” y “Colocar” la PCB en la estación “Dos”. Para salir de una zona se debe efectuar la seña “Salir”.

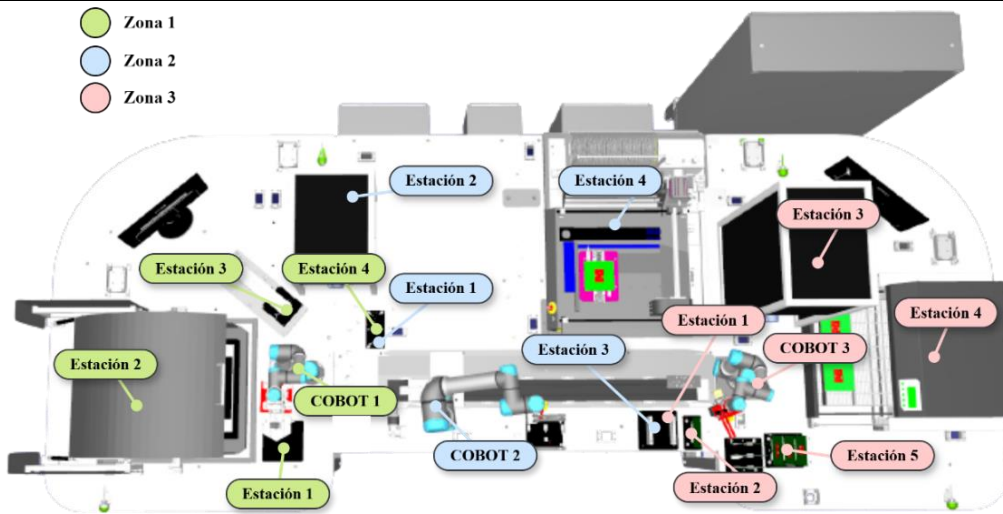


Figura 3. Celda de manufactura de PCBs.

Para evaluar la precisión de la HMI se hicieron pruebas con 10 usuarios diferentes, cada uno efectuó las señas para cada mano en repetidas ocasiones. La Figura 4 (a) y (b) corresponde a las matrices de confusión de señas para mano izquierda y derecha respectivamente. En (a) se observa que la seña “Mover” tiene la mejor precisión con 99 clasificaciones correctas, la seña “Tomar” fue clasificado erróneamente como “Acceder/Salir” en 4 ocasiones y como “Colocar” en 5 ocasiones. En general, los errores son bajos, lo que demuestra que hay un buen desempeño en la clasificación de señas de la mano izquierda. Por otro lado, en (b) se observa que las señas “Uno” y “Cinco” fueron clasificados correctamente el 100% de las veces, mientras que el resto tuvieron algunos errores de clasificación, sin embargo, aunque hay ciertas confusiones, la precisión global para el reconocimiento de señas de la mano derecha es bastante aceptable.

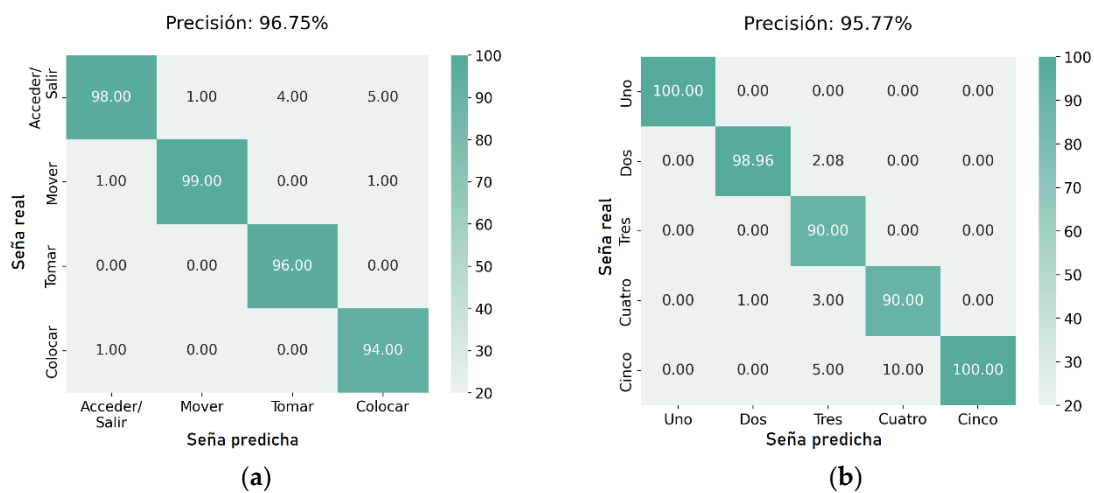


Figura 4. Matrices de confusión: (a) Desempeño de la HMI para la clasificación de señas de la mano izquierda; b) Desempeño de la HMI para la clasificación de señas de la mano derecha.

5. Conclusiones

El uso del dispositivo LMC como herramienta para el reconocimiento de señas estáticas ha demostrado ser una opción viable para la creación de una HMI ya que permitió tener una interacción más segura y natural con los cobots. La precisión del reconocimiento de señas alcanzó un 96.75% para las señas con la mano izquierda y un 95.77% para las señas con la mano derecha, lo que demuestra la



eficiencia del dispositivo. Es esencial considerar la iluminación del entorno para evitar comportamientos erráticos en el reconocimiento de las señas. En la Tabla 2 se muestran las principales ventajas y desventajas de controlar cobots mediante señas de las manos utilizando LMC.

Tabla 2. Ventajas y desventajas de controlar cobots mediante señas de las manos utilizando el LMC.

Ventajas	Desventajas
La interacción sin contacto incrementa la seguridad del usuario.	El LMC tiene un rango de visión limitado.
Permite una interacción natural e intuitiva al no requerir de conocimientos técnicos avanzados.	El reconocimiento de señas presenta Inconsistencia bajo ciertas características morfológicas de las manos.
Genera una brecha de oportunidad a personas con diferentes capacidades motrices.	Inconsistencia ante iluminación cenital.

Financiamiento: Esta investigación fue financiada por el Consejo Nacional de Ciencias, Humanidades y Tecnología de México (CONAHCYT) con el número de subvención F003-322609.

Agradecimientos: Nuestro agradecimiento al Centro de Ingeniería y Desarrollo Industrial (CIDESI) por proporcionarnos el espacio y herramientas necesarias para el desarrollo y culminación de este proyecto. Asimismo, agradecemos al equipo del Laboratorio Nacional de Investigación en Tecnologías Digitales (LANITED) por su inestimable contribución a lo largo del desarrollo del proyecto.

Declaración sobre disponibilidad de datos: Los datos que sustentan los resultados de este estudio pueden solicitarse al autor de correspondencia.

Conflictos de intereses: Los autores declaran no tener conflictos de intereses.

Referencias

- [1] M. Tölgyessy, M. Dekan, J. Rodina, y F. Duchoň, "Analysis of the Leap Motion Controller workspace for HRI gesture applications", *Appl. Sci. (Basel)*, vol. 13, núm. 2, p. 742, 2023. <https://doi.org/10.3390/app13020742>.
- [2] X. Zhang, R. Zhang, L. Chen, y X. Zhang, "Natural gesture control of a delta robot using leap motion", *J. Phys. Conf. Ser.*, vol. 1187, núm. 3, p. 032042, 2019. <https://doi.org/10.1088/1742-6596/1187/3/032042>.
- [3] C. Li, A. Fahmy, y J. Sienz, "Development of a neural network-based control system for the DLR-HIT II robot hand using leap motion", *IEEE Access*, vol. 7, pp. 136914–136923, 2019. <https://doi.org/10.1109/ACCESS.2019.2942648>.
- [4] G. Bonet, "Robot workspace sensing and control with Leap Motion Sensor," *Proceedings of the Semantic Scholar Conference*, 2013. <https://api.semanticscholar.org/CorpusID:107739948>.
- [5] O. Justinico, P. F. Cárdenas, and J. S. R. Rodríguez, "Control gestual de robot de 4 GDL con sensor Leap Motion," *Memorias*, 2015. <https://api.semanticscholar.org/CorpusID:148229251>.
- [6] J. E. López-Flores, "Control por gestos usando Leap Motion," *Trabajo de obtención de grado, Especialidad en Sistemas Embebidos, Tlaquepaque, Jalisco: ITESO*, 2016. <https://rei.iteso.mx/items/c2d4f97d-6018-4fd7-b62c-b4008dff4367>.
- [7] M. Čorňák, M. Tölgyessy, and P. Hubinský, "Innovative Collaborative Method for Interaction between a Human Operator and Robotic Manipulator Using Pointing Gestures," *Applied Sciences*, vol. 12, no. 1, p. 258, 2022. <https://doi.org/10.3390/app12010258>.
- [8] R. Nogales, F. Mayorga, and J. Vargas, "A proposal for Hand gesture control applied to the KUKA youBot using motion tracker sensors and machine learning algorithms," in *2022 17th Iberian Conference on Information Systems and Technologies (CISTI)*, 2022, pp. 1-4. <https://doi.org/10.23919/CISTI54924.2022.9820218>.
- [9] P. Neto, M. Simão, N. Mendes, M. Safeea. "Gesture-based human-robot interaction for human assistance in manufacturing". *Int J Adv Manuf Technol* 101, 119–135 (2019). <https://doi.org/10.1007/s00170-018-2788-x>.



-
- [10] S. Wang, A. Wang, M. Ran, L. Liu, Y. Peng, M. Liu, G. Su, A. Alhudhaif, F. Alenezi, N. Alnaim. "Hand gesture recognition framework using a lie group based spatio-temporal recurrent network with multiple hand-worn motion sensors". *Information Sciences*. 2022; 606: 722-741. ISSN 0020-0255. <https://doi.org/10.1016/j.ins.2022.05.085>.
- [11] R. Mariappan, P. Gayathri, P. Pushpalatha, V. Sri Sidhvi Rishik, and T. Satish, "Real Time Robotic Arm using Leap Motion Controller," *Journal of Physics: Conference Series*, vol. 2466, no. 1, p. 012023, Mar. 2023. <https://doi.org/10.1088/1742-6596/2466/1/012023>.
- [12] K. Chatterjee, M. Raju, N. Selvamuthukumar, M. Pramod, B. Krishna Kumar, A. Bandyopadhyay y S. Mallik, "HaCk: Hand Gesture Classification Using a Convolutional Neural Network and Generative Adversarial Network-Based Data Generation Model," *Information*, vol. 15, núm. 2, art. núm. 85, 2024. <https://www.mdpi.com/2078-2489/15/2/85>
- [13] H. Zhang, K. Liu, Y. Zhang y J. Lin, "TRANS-CNN-Based Gesture Recognition for mmWave Radar," *Sensors*, vol. 24, núm. 6, art. núm. 1800, 2024. <https://www.mdpi.com/1424-8220/24/6/1800>
- [14] W. Qi, S. E. Ovrur, Z. Li, A. Marzullo y R. Song, "Multi-Sensor Guided Hand Gesture Recognition for a Teleoperated Robot Using a Recurrent Neural Network," *IEEE Robotics and Automation Letters*, vol. 6, núm. 3, pp. 6039–6045, 2021. <https://doi.org/10.1109/LRA.2021.3089999>
- [15] H. Liu y Z. Liu, "A Multimodal Dynamic Hand Gesture Recognition Based on Radar–Vision Fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–15, 2023. <https://doi.org/10.1109/TIM.2023.3253906>