

Teleoperación háptica de una mano robótica para manipulación remota inteligente

Gomariz, J.^a, Bernat, L.^a, Romero, C.^a, Ubeda, A.^{a,*}, Jara, C. A.^a

^aGrupo Human Robotics, Universidad de Alicante, Crta. de San Vicente del Raspeig S/N, 03690, San Vicente del Raspeig, España.

Resumen

El control natural de robots humanoides es un área de interés en robótica porque permite trasladar la destreza y la sensibilidad humanas a sistemas robóticos que operan a distancia. En este contexto, este proyecto propone un sistema de teleoperación háptica aplicado a una mano robótica antropomórfica, pensado tanto para la manipulación remota como para el estudio de interfaces avanzadas entre personas y robots. La metodología se basa en captar el movimiento de la mano humana mediante técnicas de visión artificial, de forma que los gestos del operador se convierten en tiempo real en movimientos de las articulaciones de la mano robótica. Al mismo tiempo, la retroalimentación háptica vibrotáctil aporta información sensorial relacionada con el contacto de la mano robótica con el entorno, lo que mejora la percepción del usuario y la precisión al manipular objetos. Las pruebas realizadas muestran que el sistema funciona correctamente, lo que confirma su potencial como plataforma de investigación y desarrollo, con posibilidades de evolucionar hacia aplicaciones reales de teleoperación y sistemas de mayor complejidad.

Palabras clave: Sistemas hombre-máquina, Sistemas multisensoriales, Interacción multimodal, Robots manipuladores, Telerrobótica

Haptic teleoperation of a robotic hand for intelligent robotic manipulation

Abstract

Natural control of humanoid robots is an area of interest in robotics since it enables the transfer of human dexterity and sensitivity to robotic systems operating at a distance. In this context, this project proposes a haptic teleoperation system applied to an anthropomorphic robotic hand, intended both for remote manipulation and for the study of advanced human-robot interfaces. The methodology is based on capturing the movement of the human hand using computer vision techniques, so that the operator's gestures are translated in real time into joint movements of the robotic hand. At the same time, vibrotactile haptic feedback provides sensory information related to the contact between the robotic hand and the environment, improving the user's perception and precision when manipulating objects. The experimental results show that the system operates correctly, confirming its potential as a research and development platform, with the possibility of evolving toward real teleoperation applications and more complex systems.

Keywords: Human-machine systems, Multi sensor systems, Multi-modal interaction, Robots manipulators, Telerobotics

1. Introducción

La importancia de los sistemas de teleoperación háptica en manipulación reside en que permiten interactuar de forma precisa y segura con entornos remotos, transmitiendo al operador información motora y sensorial relevante mientras realiza una tarea (7). En la práctica, la teleoperación robótica ya ha demostrado su utilidad en aplicaciones reales. Un ejemplo claro son los robots utilizados en la desactivación de explosivos, donde los manipuladores teleoperados permiten trabajar a distancia en

entornos extremadamente peligrosos. Sin embargo, estos sistemas suelen emplear controladores tipo joystick, lo que reduce la precisión y naturalidad del control. Algo similar ocurre en el ámbito médico: plataformas quirúrgicas como el da Vinci Surgical System (1) han supuesto un avance decisivo en la cirugía mínimamente invasiva al permitir operar a distancia. Aun así, en muchos casos la retroalimentación háptica es limitada o inexistente, lo que disminuye la sensación de inmersión del cirujano en el entorno remoto.

Para mejorar esa naturalidad en la interacción, los guantes

*Autor para correspondencia: andres.ubeda@ua.es

hápticos han ido ganando protagonismo como interfaces de entrada y salida en sistemas de teleoperación y realidad virtual (10). Soluciones comerciales como los guantes SenseGlove (2) permiten capturar el movimiento de la mano e incorporar retroalimentación háptica de fuerza, y se utilizan ampliamente en simulación, formación y entornos virtuales. En esta misma línea, los guantes TactGlove DK2 de bHaptics (3), empleados en este proyecto, ofrecen retroalimentación vibrotáctil que facilita la transmisión de información sensorial de forma no invasiva, lo que los hace especialmente adecuados para entornos experimentales y educativos.

En el ámbito de la manipulación robótica, las manos antropomórficas se consideran algunos de los dispositivos más complejos y versátiles dentro de la robótica de intervención. La mano robótica RH8D de Seed Robotics (?), utilizada en este trabajo, destaca por su diseño antropomórfico y su elevado número de grados de libertad, lo que la convierte en una plataforma muy apropiada para estudiar estrategias de control natural y teleoperación basada en gestos humanos.

Además, existen proyectos de investigación y desarrollos comerciales que combinan manos robóticas antropomórficas con sistemas avanzados de teleoperación. El proyecto HandyBot, desarrollado por INESCOP (Instituto Tecnológico del Calzado y Conexas) (5), permite teleoperar una mano robótica integrada en un brazo manipulador. Por su parte, las soluciones de Shadow Robot Company (?) son uno de los principales referentes en teleoperación háptica de manos robóticas, ya que ofrecen sistemas capaces de reproducir con gran fidelidad los movimientos de la mano humana, aunque con una elevada complejidad técnica y un coste considerable.

Desde esta perspectiva, el proyecto desarrollado integra la mano robótica RH8D y los guantes hápticos TactGlove DK2 para crear un prototipo de teleoperación háptica que resulta intuitivo, fácil de utilizar y capaz de funcionar en tiempo real desde cualquier parte del mundo.

2. Desarrollo e implementación

2.1. Herramientas

Este trabajo se apoya en tecnologías ampliamente utilizadas en el ámbito de la robótica y la ingeniería. En concreto:

- **ROS:** Robot Operating System, empleado como herramienta para gestionar los nodos del sistema y facilitar la comunicación con la mano robótica.
- **VPN:** Red Privada Virtual, utilizada para garantizar una conexión segura y rápida entre los distintos subsistemas. En este caso, la plataforma empleada es Tailscale.
- **MediaPipe:** modelo de estimación de pose basado en inteligencia artificial, utilizado para el reconocimiento y procesamiento de gestos en tiempo real.

2.2. Arquitectura general

La arquitectura propuesta se compone de dos elementos principales:

- **Cliente:** un ordenador con Windows que se encarga de capturar los gestos de la mano y de controlar los guantes hápticos Tactglove DK2.

- **Servidor:** un ordenador con Ubuntu 20.04 conectado directamente a la mano robótica RH8D.

Ambos equipos se comunican a través de una VPN, lo que garantiza una conexión rápida y segura, independientemente de la distancia o la ubicación física entre ellos.

En la Figura 1 se presenta un esquema general del funcionamiento completo del sistema:

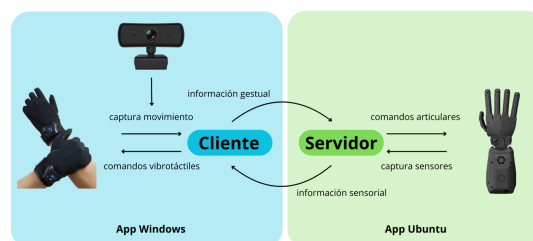


Figura 1: Esquema general de la arquitectura del sistema.

El diagrama posee dos líneas de información distintas:

Control de la mano

1. Una cámara integrada en el equipo cliente captura el movimiento de la mano del operador.
2. El cliente procesa esa información gestual y la envía al servidor.
3. En el servidor, los datos recibidos se transforman en comandos articulares que se transmiten a la mano robótica.

Retroalimentación sensorial

1. Mientras se ejecuta el movimiento, la mano robótica recoge información del entorno a través de sus sensores.
2. El servidor convierte esas señales en información adecuada para la retroalimentación y la envía al cliente.
3. El cliente traduce la información recibida en comandos de estimulación vibrotáctil que se transmiten a los guantes hápticos.

Desde el punto de vista funcional, este diseño se ha organizado en varios subsistemas, cada uno responsable de una etapa concreta del proceso de teleoperación. Trabajando de forma coordinada, todos ellos permiten que la teleoperación háptica de la mano robótica se realice en tiempo real.

2.3. Visión artificial y captura de gestos

El módulo de visión artificial permite estimar la pose de la mano humana usando una única cámara RGB. Para ello, se emplea MediaPipe Hands, que detecta 21 puntos característicos (landmarks) de la mano, incluyendo muñeca y articulaciones de los dedos (Figura 2):

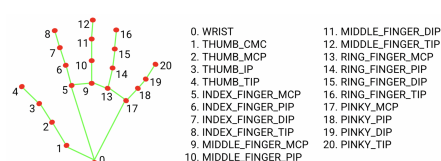


Figura 2: Landmarks de MediaPipe

Cada punto se representa como:

$$\mathbf{p}_i = (x_i, y_i, z_i), \quad (1)$$

donde x_i, y_i es la coordenada del punto en la imagen y z_i representa una profundidad relativa estimada por el modelo. A partir de estos puntos se puede calcular:

- El grado de apertura o cierre de cada dedo.
- La orientación de la muñeca.

2.3.1. Filtrado de la señal

La detección puede presentar pequeñas oscilaciones debido a iluminación, ruido o movimientos rápidos. Para suavizar la señal se aplica un filtro exponencial (EMA):

$$\hat{\mathbf{p}}_i(t) = \alpha \mathbf{p}_i(t) + (1 - \alpha) \hat{\mathbf{p}}_i(t - 1), \quad (2)$$

donde disminuir el factor $\alpha \in (0, 1)$ aumenta el suavizado de la señal. Este filtro reduce vibraciones bruscas sin introducir un retardo significativo.

2.3.2. Cálculo de apertura y cierre de los dedos

Para estimar la flexión de cada dedo se calculan ángulos entre tres puntos consecutivos utilizando el producto escalar:

$$\theta(a, b, c) = \arccos \left(\frac{(\mathbf{p}_a - \mathbf{p}_b) \cdot (\mathbf{p}_c - \mathbf{p}_b)}{\|\mathbf{p}_a - \mathbf{p}_b\| \|\mathbf{p}_c - \mathbf{p}_b\|} \right) \quad (3)$$

En los dedos índice, corazón, anular y meñique, se suma el ángulo que forman los puntos MCP, PIP y DIP con el que forman los puntos PIP, DIP y TIP del dedo:

$$F_d = \theta(\text{MCP}, \text{PIP}, \text{DIP}) + \theta(\text{PIP}, \text{DIP}, \text{TIP}) \quad (4)$$

En el pulgar se realiza una suma equivalente adaptada a su cadena articular.

2.3.3. Cálculo de la orientación de la muñeca

A pesar de que se utiliza una cámara 2D, MediaPipe proporciona una estimación relativa de profundidad que permite aproximar la orientación de la palma. Para ello, se construye un sistema de referencia local de la mano usando tres puntos clave:

- Muñeca: \mathbf{p}_w
- Base del índice: \mathbf{p}_i
- Base del meñique: \mathbf{p}_p

Primero, se define un vector que apunta desde la muñeca hacia el índice:

$$\mathbf{x}_h = \frac{\mathbf{p}_i - \mathbf{p}_w}{\|\mathbf{p}_i - \mathbf{p}_w\|} \quad (5)$$

Después, se define un segundo vector hacia el meñique:

$$\mathbf{y}_{raw} = \mathbf{p}_p - \mathbf{p}_w \quad (6)$$

Con estos dos vectores se calcula la normal de la palma mediante el producto vectorial:

$$\mathbf{z}_h = \frac{\mathbf{x}_h \times \mathbf{y}_{raw}}{\|\mathbf{x}_h \times \mathbf{y}_{raw}\|} \quad (7)$$

Finalmente, se obtiene el tercer eje ortogonal:

$$\mathbf{y}_h = \mathbf{z}_h \times \mathbf{x}_h \quad (8)$$

De este modo se construye un sistema de referencia local orthonormal $(\mathbf{x}_h, \mathbf{y}_h, \mathbf{z}_h)$ asociado a la mano. A partir de este, se calculan los tres ángulos principales de la muñeca:

- **Aducción:** desviación lateral.

$$\mathbf{yaw} = \arctan(x_{h,y}, x_{h,x}) \quad (9)$$

- **Flexión:** inclinación hacia delante o hacia atrás.

$$\mathbf{pitch} = \arctan \left(-x_{h,z}, \sqrt{x_{h,x}^2 + x_{h,y}^2} \right) \quad (10)$$

- **Roll:** rotación sobre el eje longitudinal.

$$\mathbf{roll} = \arctan(y_{h,z}, z_{h,z}) \quad (11)$$

2.3.4. Auto-calibración

Cada persona tiene una morfología y un rango de movimiento distintos. Para adaptarse automáticamente a estas diferencias, el sistema registra valores mínimos y máximos durante una fase inicial de calibración. Este proceso se realiza únicamente al comienzo.

A partir de ese momento, durante el movimiento normal de la mano, cada señal se normaliza al intervalo $[0,1]$. De esta forma, los valores que se envían al servidor son coherentes, independientemente del usuario o de la distancia a la cámara. Esta normalización se lleva a cabo en cada iteración del sistema. Además, se aplica una pequeña corrección no lineal con el objetivo de mejorar la sensibilidad en determinadas zonas del movimiento.

2.3.5. Procesado HSV para uso con guantes

MediaPipe está pensado principalmente para detectar manos desnudas. Cuando se utilizan guantes, la estimación de la pose puede verse afectada. Para conseguir un funcionamiento correcto con el guante puesto, se aplica un desplazamiento del tono en el espacio HSV.

Este valor se obtuvo de forma experimental, probando distintas configuraciones hasta encontrar una que ofreciera una detección estable con el guante, pero que también funcionara correctamente con la mano desnuda. Tras varias pruebas, se estableció como adecuado un valor de 15 (Figura 3).

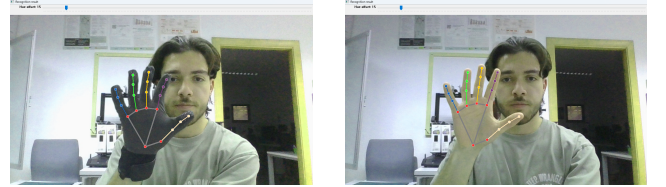


Figura 3: Comparación de la detección de landmarks con guante (izquierda) y sin guante (derecha).

2.4. Teleoperación y control de la mano robótica

Este subsistema se ejecuta en el servidor bajo ROS. Su función es recibir los valores enviados por el cliente, convertirlos en posiciones reales de las articulaciones de la mano RH8D y asegurar que el movimiento sea estable y seguro.

2.4.1. Conversión a posiciones articulares reales

Para transformar los datos de flexión de cada dedo en posiciones reales de las articulaciones se utiliza un escalado lineal:

$$q = q_{min} + u \cdot (q_{max} - q_{min}) \quad (12)$$

donde u es el valor recibido, q_{min} y q_{max} son los límites de cada articulación, y q es la posición final que se envía al robot. Los valores q_{min} y q_{max} se determinaron experimentalmente, hasta conseguir un movimiento del robot adaptado a su rango real útil.

2.4.2. Filtro de Kalman

Aunque la visión artificial ya incorpora filtrado, la señal puede seguir teniendo ruido o cambios bruscos. Para evitar movimientos poco naturales y oscilatorios, cada articulación se filtra utilizando un filtro de Kalman.

Se utiliza un modelo básico donde el estado, la evolución y la medición se basan en las ecuaciones (13), (14) y (15), respectivamente:

$$x_k = q_k \quad (13)$$

$$x_k = x_{k-1} + w_k \quad (14)$$

$$z_k = x_k + v_k \quad (15)$$

donde w_k representa pequeñas variaciones del sistema y v_k representa ruido en la medición. El filtro funciona en dos pasos:

1. Predicción

$$\hat{x}_{k|k-1} = \hat{x}_{k-1} \quad (16)$$

2. Corrección

$$K_k = \frac{P_{k|k-1}}{P_{k|k-1} + R} \quad (17)$$

$$\hat{x}_k = \hat{x}_{k|k-1} + K_k (z_k - \hat{x}_{k|k-1}) \quad (18)$$

En términos sencillos, reduce vibraciones y evita movimientos bruscos manteniendo una buena capacidad de respuesta.

2.4.3. Postura segura

En los casos de pérdida de conexión con el cliente, de no detección de la mano humana u otro tipo de errores, se protege al sistema con una postura segura preestablecida. Esta pose consiste en colocar todas las articulaciones de la mano robótica en un rango intermedio estable.

2.5. Retroalimentación háptica

La retroalimentación háptica tiene como objetivo transmitir al operador información sobre el contacto entre la mano robótica y el entorno. Para ello, se convierte la señal de fuerza medida en las yemas de los dedos de la RH8D en estímulos vibrotáctiles localizados en los guantes del usuario.

2.5.1. Procesamiento de la señal de fuerza

Cada dedo proporciona el componente de la fuerza en tres ejes (f_x, f_y, f_z). Para obtener una medida independiente de la dirección del contacto, se calcula la magnitud total de la fuerza:

$$m = \sqrt{f_x^2 + f_y^2 + f_z^2} \quad (19)$$

2.5.2. Compensación de ruido y filtrado de la señal

Incluso cuando no existe contacto real, los sensores pueden introducir ruido. Para evitar estas vibraciones en el guante, se estima un nivel de base o baseline:

$$b_k = (1 - \alpha_b) b_{k-1} + \alpha_b m_k \quad (20)$$

Posteriormente, se resta este nivel de base a la señal medida:

$$m_c = \text{máx}(0, m - b_k) \quad (21)$$

Esto permite que únicamente se considere contacto cuando la fuerza supera claramente el nivel de ruido del sensor. Además, se aplica un filtrado exponencial (EMA) para reducir aún más las oscilaciones.

2.5.3. Conversión a intensidad vibrotáctil

El contacto se considera válido si la señal supera un cierto umbral de activación T_{on} . Por tanto, la magnitud efectiva se calcula como:

$$m_{eff} = \frac{m - T_{on}}{1 - T_{on}} \quad (22)$$

donde se asume que la señal máxima está próxima a 1. La intensidad vibrotáctil aplicada al actuador corresponde con la magnitud efectiva. De esta forma:

- Si $m = T_{on} \rightarrow I = 0$
- Si $m = 1 \rightarrow I = 1$
- Valores intermedios generan vibraciones proporcionales.

2.5.4. Justificación del uso de retroalimentación vibrotáctil

En este sistema se ha elegido la retroalimentación vibrotáctil en lugar de la realimentación de fuerza activa por varios motivos:

- Los guantes vibrotáctiles son más ligeros y cómodos de usar.
- Su coste es menor y su integración en el sistema resulta más sencilla.

Aunque la realimentación de fuerza puede proporcionar una sensación más realista, la opción vibrotáctil es suficiente para tareas de manipulación básica y para entornos experimentales.

2.6. Interfaz de usuario

El sistema dispone de dos aplicaciones independientes: una para el servidor (Ubuntu) y otra para el cliente (Windows). Cada una está diseñada para gestionar el funcionamiento de la teleoperación según las necesidades de su entorno.

2.6.1. Aplicación del servidor (Ubuntu)

Al iniciarse, la aplicación solicita la contraseña del usuario para poder ejecutar comandos con privilegios de superusuario, necesarios para gestionar procesos relacionados con ROS y la comunicación con la mano robótica. Una vez validada, se muestra una interfaz sencilla donde se indica el estado general del sistema, los nodos activos y otra información básica de funcionamiento. También incorpora un botón de parada de emergencia. Cuando se activa, la mano robótica se detiene inmediatamente y mantiene su posición actual hasta que el sistema vuelve a activarse.

2.6.2. Aplicación del cliente (Windows)

La aplicación del cliente está organizada en varias pantallas que guían al operador durante la configuración y el uso del sistema.

Ventana inicial

En esta primera pantalla se muestran unas instrucciones básicas y se solicita la configuración de la dirección IP y el puerto de comunicación. En esta fase se establece la conexión con el servidor antes de comenzar la teleoperación.

Ventana de calibración

Al seleccionar esta opción, el usuario debe abrir y cerrar el puño varias veces durante unos segundos. Con este procedimiento se realiza la auto-calibración descrita en el apartado 3.2.4.:

Ventana principal

La ventana principal se divide en dos áreas (Figura 4):

- **Panel izquierdo:** muestra la imagen capturada por la cámara del operador junto con los *landmarks* detectados por MediaPipe. Esto permite comprobar visualmente que la detección de la mano se está realizando correctamente.
- **Panel derecho:** presenta un modelo 3D de la mano robótica desarrollado en Unity. Este modelo está sincronizado con la mano real y reproduce en tiempo real los movimientos que ejecuta el robot.

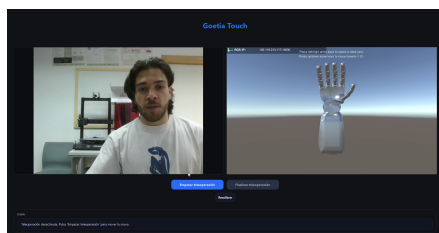


Figura 4: Ventana principal

Gracias a esta disposición, el operador puede conocer el estado de la mano robótica sin necesidad de encontrarse físicamente en el entorno remoto. En esta misma pantalla se incluyen botones para iniciar y detener la teleoperación, realizar una nueva calibración y consultar una consola de estado donde se muestran mensajes informativos. Por razones de seguridad, al acceder a esta ventana la teleoperación permanece desactivada hasta que el usuario la activa manualmente.

3. Resultados

El sistema se evaluó de forma experimental, analizando su comportamiento en tiempo real con toda la arquitectura en funcionamiento. Durante las pruebas se comprobó que la metodología desarrollada permite una teleoperación fluida y estable. Dentro del contexto en el que se ha llevado a cabo el proyecto, no se observaron retardos ni oscilaciones significativas.

Para validar el sistema de teleoperación, se llevaron a cabo distintos ejemplos:

- Apertura y cierre de la mano (Figura 5).



Figura 5: Apertura y cierre de la mano.

- Movimientos de muñeca: rotación lateral simulando un saludo (Figura 6).



Figura 6: Movimientos de muñeca.

- Gestos: como levantar dedos sueltos o realizar el gesto de la "paz" (Figura 7).



Figura 7: Gestos.

- Control parcial del cierre de los dedos: demostrando que el sistema no funciona solo en estados extremos (abierto/cerrado), sino que permite un control continuo en todo el rango de movimiento (Figura 8).



Figura 8: Control continuo.

- Manipulación de objetos, como coger una pelota (Figura 9).



Figura 9: Manipulación de una pelota.

En lo que respecta a la retroalimentación háptica, se comprobó que la vibración de los guantes se corresponde de forma directa con el contacto detectado por los sensores de la mano robótica. Esta sensación se aprecia con claridad en tareas de manipulación, como al coger una pelota, donde el operador puede notar el momento en que se produce el contacto con el objeto. Además, se verificó que cada sensor funciona de manera independiente: si se activa únicamente el sensor de un dedo, solo vibra el actuador correspondiente en el guante.

4. Conclusiones

Los resultados experimentales muestran que el sistema alcanza su objetivo principal: hacer posible una teleoperación háptica en tiempo real, con movimientos fluidos y retroalimentación sensorial. Esto confirma que tanto la arquitectura propuesta como la metodología desarrollada son adecuadas. Como posibles líneas de mejora en el futuro, se plantean las siguientes opciones:

- Sustituir los guantes vibrotáctiles por guantes con realimentación de fuerza activa, con el fin de ofrecer una sensación más realista.
- Integrar el sistema en entornos de realidad virtual.
- Instalar la mano robótica en el extremo de un brazo manipulador y ampliar el control para incluir también el movimiento del brazo.
- Incorporar una segunda mano robótica para simular ambos brazos de una persona.

Estas mejoras permitirían llevar el sistema un paso más allá, acercándolo a aplicaciones más complejas y realistas.

Agradecimientos

Este trabajo forma parte del proyecto MYOREHAB (PCI2023-143405) financiado por el Ministerio de Ciencia, Innovación y Universidades (MCIN) y la Agencia Estatal de Investigación (AEI) [MCIN/AEI/10.13039/501100011033]

Referencias

- [1] Intuitive Surgical, *da Vinci Surgical System*. Disponible en: <https://www.intuitive.com/en-us/products-and-services/da-vinci>, 2024.
- [2] SenseGlove, *SenseGlove*. Disponible en: <https://www.senseglove.com/>, 2024.
- [3] bHaptics, *TactGlove DK2*. Disponible en: <https://www.bhaptics.com/es/tactsuit/tactglove-dk2/>, 2024.
- [4] Seed Robotics, *RH8D Adult Robot Hand*. Disponible en: <https://www.seedrobotics.com/rh8d-adult-robot-hand>, 2024.
- [5] INESCOP, *HandyBot – Teleoperation of a robotic hand*. Disponible en: <https://www.inescop.es/es/i-d-i/proyectos-i-d-i/proyectos-i-d-i-ivace/54-2022/628-handybot>, 2022.
- [6] Shadow Robot Company, *Teleoperation Solutions*. Disponible en: <https://shadowrobot.com/teleoperation/>, 2024.
- [7] El Rassi, I.; El Rassi, J.-M., *A review of haptic feedback in tele-operated robotic surgery*. Journal of Medical Engineering & Technology, 2020. DOI: 10.1080/03091902.2020.1772391.
- [8] Li, C.; Yang, C.; Wan, J.; Annamalai, A. S. K.; Cangelosi, A., *Teleoperation control of Baxter robot using Kalman filter-based sensor fusion*. Systems Science & Control Engineering, 2017. DOI: 10.1080/21642583.2017.1300109.
- [9] Martínez García, A. M.; Sánchez Martínez, D.; Úbeda Castellanos, A.; Jara Bravo, C. A.; Morell Giménez, V., *Control y virtualización de una mano robótica RH8D mediante ROS*. Jornadas de Automática, 2025. DOI: 10.17979/ja-cea.2025.46.12233.
- [10] Papakonstantinou, C.; Giannakos, K.; Kokkonis, G.; Papadopoulou, M. S., *The effect of tactile feedback on the manipulation of a remote robotic arm via a haptic glove*. Electronics, 2025. DOI: 10.3390/electronics14244964.
- [11] Google, *MediaPipe Hands*. Disponible en: https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker?hl=es-419, 2024.