

Control inteligente de robots mediante gestos de la mano

La Smart Manufacturing, que forma parte de la industria 4.0, optimiza la manufactura mediante tecnologías avanzadas como inteligencia artificial, internet de las cosas —concepto que se explica más adelante— y la ciencia de datos. Estas tecnologías crean un entorno de producción eficiente y adaptable, conectando en tiempo real máquinas, sistemas y personas. Este artículo presenta un sistema ciberfísico para controlar robots colaborativos dentro de una **celda de manufactura** mediante reconocimiento de gestos.

Celda de manufactura

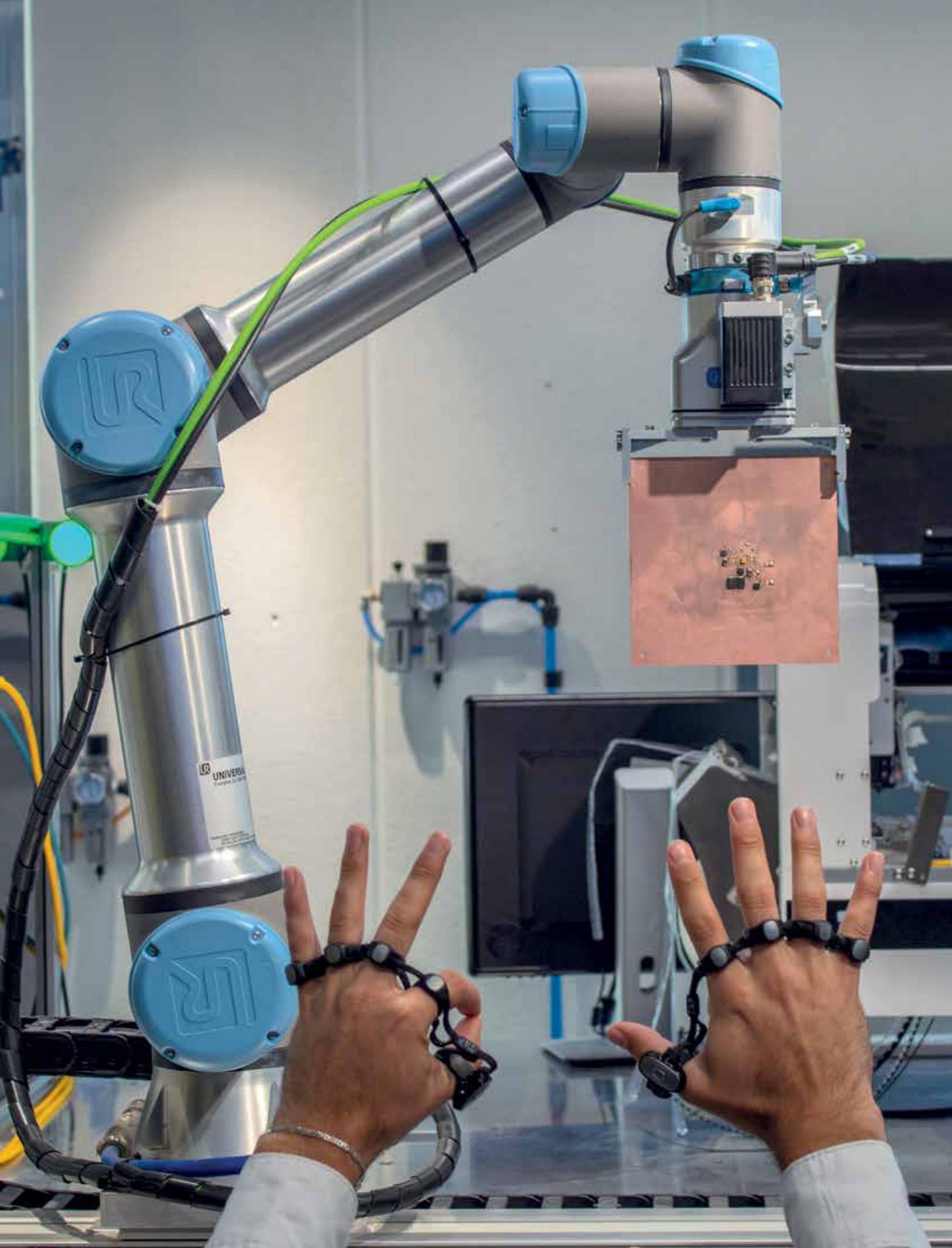
Sistema de producción automatizado y adaptable que utiliza tecnología avanzada.

Introducción

Imaginemos un futuro donde el ingenio humano y la tecnología convergen, uniendo lo tangible con lo intangible para superar las barreras físicas. Las manos, con sus gestos precisos, se convierten en el puente entre la mente y el mundo, facilitando no sólo nuestras tareas cotidianas, sino también el impulso vital de la productividad. Sin embargo, ¿qué ocurre cuando esas manos encuentran límites que les impiden moverse con libertad?

En las fábricas, donde el corazón de la industria late al compás de la colaboración entre operario y máquina, la exclusión de personas con capacidades motrices limitadas no sólo es una sombra sobre la inclusión social, sino una pérdida del valioso potencial humano. La falta de diversidad no sólo afecta a la persona, sino al dinamismo y la creatividad que florecen en la colaboración.

Frente a este desafío, la ciencia nos ofrece una respuesta: tecnologías habilitadoras que amplían horizontes. Este artículo explora cómo el reconocimiento de gestos, particularmente el movimiento de una mano, puede revolucionar los sistemas de producción, permitiendo que quienes enfrentan barreras físicas puedan controlar máquinas con la elegancia de un simple movimiento. Estas innovaciones no sólo impulsan la eficiencia de la industria 4.0, sino que tejen un futuro más inclusivo, donde el simple gesto humano transforma el mundo industrial.



Internet de las cosas

Red de objetos físicos ("cosas") que llevan incorporados sensores, con *software* para intercambiar datos a través de internet.

Sensor

Dispositivo que detecta y mide cambios en el entorno, como temperatura, presión o movimiento.

Acelerómetros

Dispositivos que miden la aceleración de un objeto.

Industria 4.0 y espacios ciberfísicos

La era de la industria 4.0 está marcada por la integración de máquinas inteligentes y sistemas que se comunican entre sí y con los humanos a través de diversas infraestructuras, como el **internet de las cosas** (IoT, por sus siglas en inglés). En el corazón de esta transformación están los espacios ciberfísicos: lugares donde el mundo digital y el físico se conectan. Aquí, los **sensores** actúan como puentes, recolectando datos del mundo real, enviándolos al mundo digital para su análisis y retroalimentando al mundo físico. Un ejemplo es el Tap Strap 2, un dispositivo que se coloca en la mano y que utiliza cinco sensores especiales llamados **acelerómetros**. Estos sensores permiten detectar cómo se mueve y gira cada dedo, para que así el dispositivo interprete gestos o acciones de la mano. Cuando aplicamos técnicas de inteligencia artificial, los datos de este dispositivo pueden ser utilizados para reconocer los movimientos de la mano y luego traducirlos en comandos específicos que una máquina pueda entender. De este modo, con sólo un gesto, podríamos hacer que dicha máquina ejecute una tarea específica o una secuencia de acciones.

La comunicación entre máquinas es fundamental en entornos industriales modernos, como las fábricas avanzadas (inteligentes), donde es importante que todo funcione de manera precisa y coordinada para garantizar la eficiencia y la productividad. El desafío radica en que cada máquina se comunica con

su propio lenguaje (protocolo de comunicación), que puede variar considerablemente, creando así barreras significativas para la integración automática. Para superar estos obstáculos, existen protocolos de comunicación avanzados como el MQTT (Message Queuing Telemetry Transport), que facilita que los dispositivos y máquinas se comuniquen entre sí de manera inalámbrica a través de mensajes. MQTT funciona con una arquitectura que incluye una computadora principal (conocida como servidor) a la que llamaremos *broker* y varios dispositivos conocidos como clientes. Los mensajes entre los clientes se agrupan en categorías específicas llamadas "tópicos", y cada mensaje está relacionado con una acción específica. El *broker* actúa como intermediario central que gestiona la transmisión de mensajes entre los clientes. Los clientes pueden enviar o recibir mensajes en ciertos "tópicos". De esta manera, cuando un cliente envía un mensaje en un "tópico", el *broker* se encarga de reenviar el mensaje a todos los clientes.

Redes neuronales artificiales

En el amplio espectro de la industria 4.0, las redes neuronales artificiales (RNA) se destacan por su capacidad de imitar la manera en que el cerebro humano procesa información, lo que les permite realizar tareas complejas de reconocimiento de patrones y toma de decisiones. Existen diferentes tipos de redes neuronales, algunas más complejas que otras. Una RNA puede recibir datos de una amplia variedad de sensores como: cámaras, temperatura, humedad, movimiento y muchos otros. Los datos que estos sensores generan son procesados por "neuronas" artificiales, las cuales son un modelo matemático inspirado en las neuronas del cerebro humano. Su función es recibir datos, procesarlos mediante operaciones matemáticas simples, y generar una salida que puede ser usada para tomar decisiones. Al conjunto de neuronas se les conoce como capas, las cuales son capaces de reconocer tendencias (patrones). Por ejemplo, una RNA puede ser entrenada para que, a partir de la información que transmite el Tap Strap 2, sea capaz de reconocer gestos de la mano. Para que una RNA pueda dar respuestas más precisas, es impor-



tante usar una gran cantidad de datos mientras se le enseña cómo resolver un problema.

Caso de estudio: interfaz humano-robot

En el Laboratorio Nacional de Investigación en Tecnologías Digitales (LANITED), alojado dentro de las instalaciones del Centro de Ingeniería y Desarrollo Industrial (CIDESI) en Querétaro, se desarrollan tecnologías innovadoras asociadas a la industria 4.0. Uno de los proyectos insignia del LANITED es el desarrollo de una celda de manufactura de placas de circuito impreso (conocidas como PCB por sus siglas en inglés), que son placas utilizadas para conectar y soportar componentes electrónicos.

La celda de manufactura tiene tres robots colaborativos (conocidos como cobots) que se encargan de transportar la placa a lo largo de toda la línea de ensamble. Estos robots están diseñados para trabajar colaborativamente con humanos, por lo que son más seguros en comparación con los robots industriales tradicionales. También se compone de

diferentes estaciones, que pueden ser desde estantes (o racks) donde se coloca la materia prima (placas de cobre), hasta maquinaria que se encarga de realizar las diferentes etapas de fabricación de las PCB, como el grabado de pistas, colocación de pasta de soldadura, colocación de componentes sobre la superficie de la placa y la fijación permanente de éstos mediante el calentamiento a altas temperaturas (la **Figura 1** esquematiza la celda y lo antes explicado).

Tradicionalmente, todos los componentes que integran una celda de manufactura son manipulados por medio de una serie de instrucciones programadas previamente por un técnico. Como alternativa al control tradicional, se propone un sistema ciberfísico basado en el uso de un dispositivo Tap Strap 2 para establecer una interfaz (interacción) humano-robot por medio del reconocimiento de gestos de la mano.

El desarrollo de este sistema ciberfísico se realizó en cinco fases. Primero, se definió un “diccionario de gestos”, cada uno vinculado a una acción específica dentro de la celda. Posteriormente, se realizaron

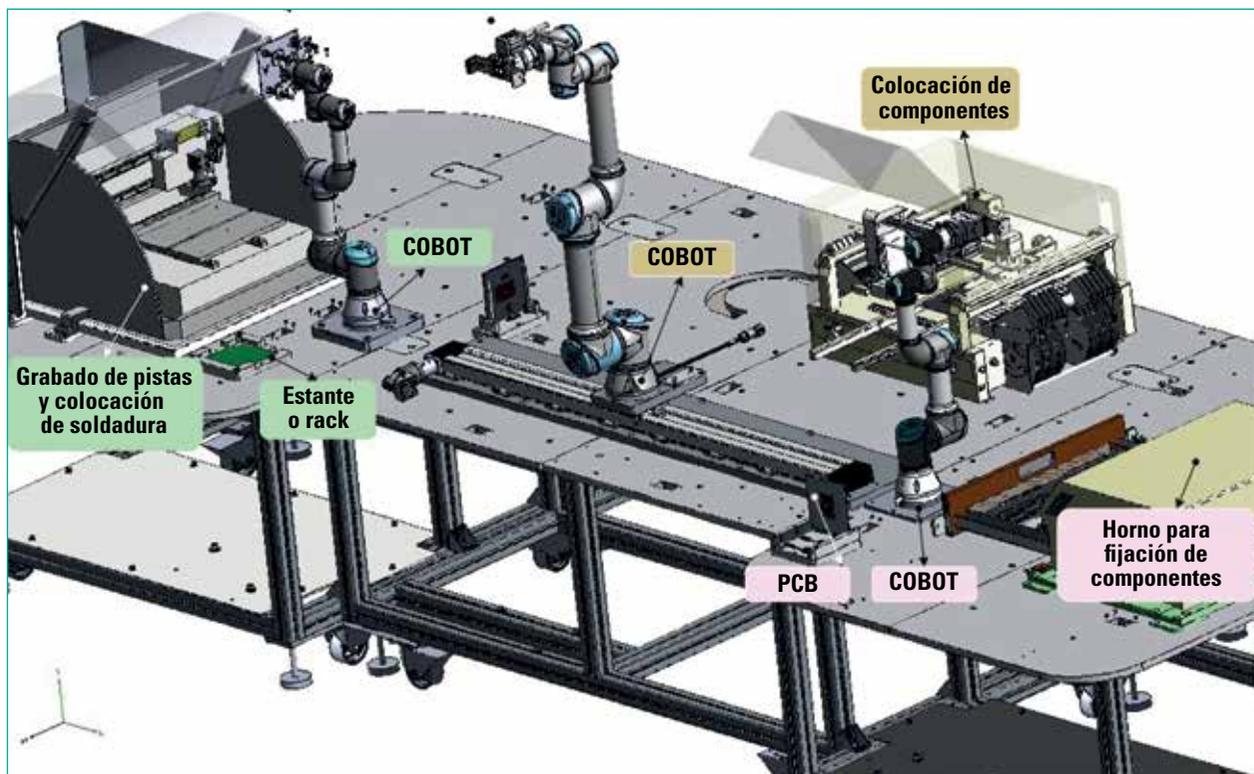


Figura 1. Celda de manufactura de PCB instalada en LANITED. Imagen elaborada por los autores.

Receptor bluetooth
Tecnología inalámbrica para la recepción de datos a corta distancia.

RSSI
Received Signal Strength Indicator, medida de la potencia de una señal recibida.

Frecuencia
Número de repeticiones de un evento por unidad de tiempo.

Archivo txt
Archivo de texto plano sin formato especial.

pruebas experimentales para conocer la calidad de la señal del dispositivo Tap Strap 2. A continuación, se adquirieron los datos transmitidos por el Tap Strap 2 con un **receptor bluetooth** integrado en un dispositivo IoT. Entonces se diseñó y configuró una RNA para el reconocimiento de los gestos de la mano. Y finalmente, se definió la lógica de comunicación entre el Tap Strap 2, el dispositivo IoT, el *broker* y los componentes de la celda.

Diccionario de gestos

La estrategia para la definición del diccionario se basó en la asignación de un gesto específico para cada componente de la celda de manufactura, así como también para cada una de las acciones que éstos son capaces de realizar (mover, tomar, colocar, abrir y cerrar). Para los componentes, se asignaron gestos alusivos al orden en que participan de acuerdo con la secuencia de producción. La **Figura 2** ilustra los gestos definidos. Por ejemplo, el primer componente de la línea de producción se identifica mediante el índice extendido, mientras que el segundo componente se identifica con los dedos índice y medio extendidos. En cuanto a las acciones, cada una se asoció con un gesto que refleja el movimiento o proceso que desencadena; por ejemplo, el gesto “tomar” indica que el cobot debe tomar la PCB. Esta estrategia fue diseñada para garantizar una curva de aprendiza-

je simple para cualquier persona que interactúe con la celda de manufactura.

Caracterización de señal

Se realizaron experimentos para caracterizar la intensidad de la **potencia de la señal recibida** (RSSI) y la **frecuencia** de los datos transmitidos por el Tap Strap 2 en función de la distancia respecto al receptor. Los resultados demostraron que el dispositivo es muy estable. Esto se evidenció por la mínima variación en la intensidad de la señal y la frecuencia de transmisión de datos, incluso al incrementar la distancia entre el dispositivo y el receptor. La **Figura 3** muestra gráficamente los resultados obtenidos.

Adquisición de datos

Se desarrolló un sistema para la adquisición de los datos del Tap Strap 2 y se integró en un equipo compacto y potente (mini PC NUC), con capacidades de cómputo de alto rendimiento, configurado para trabajar como un dispositivo IoT. El sistema establece una conexión mediante *bluetooth* entre el Tap Strap 2 (transmisor) y el dispositivo IoT (receptor), recolectando 20 000 paquetes de datos, cada uno compuesto por 15 características para cada gesto. Los datos fueron almacenados en un **archivo txt**. Después se normalizaron los datos; es decir, se limitó su varia-



Figura 2. Diccionario de gestos. Imagen elaborada por los autores.

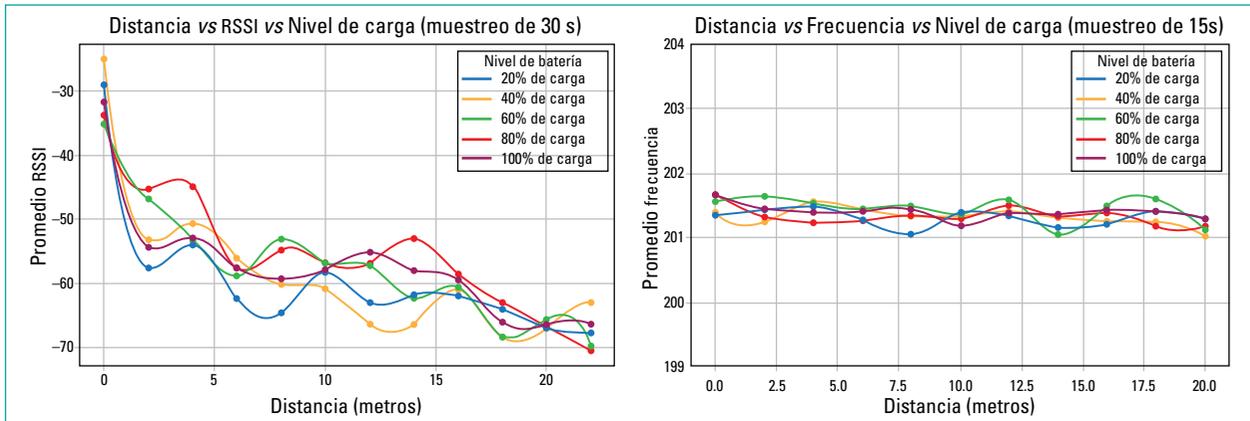


Figura 3. Intensidad de la señal y frecuencia de datos del Tap Strap 2. Gráficas elaboradas por los autores.

bilidad para que estuvieran dentro de un rango estándar entre 0 y 1. Con esto se asegura que la etapa de entrenamiento sea más rápida y precisa al acotar la escala de los datos.

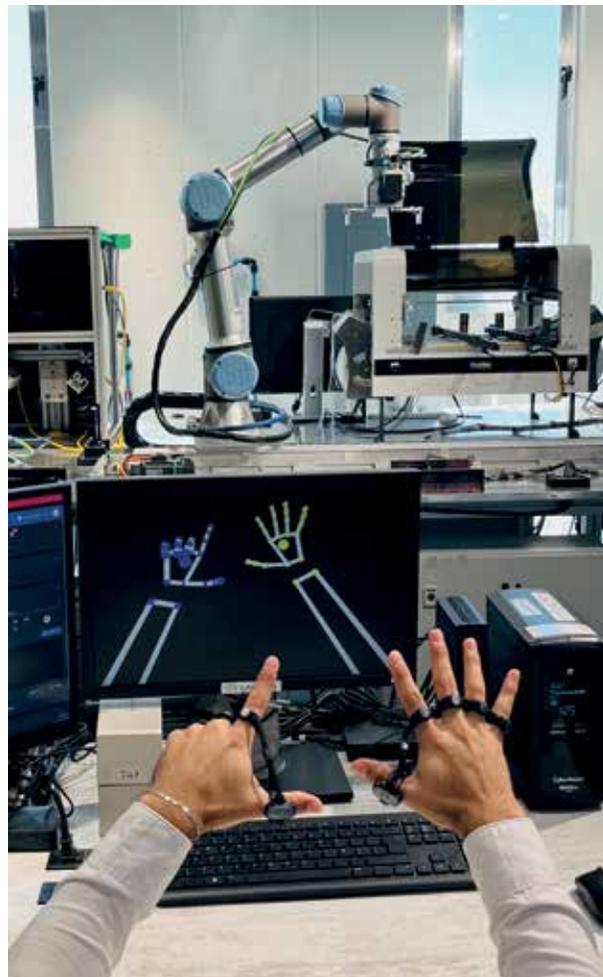
Inteligencia artificial aplicada

Se diseñó, configuró e implementó un modelo de RNA de memoria a corto-largo plazo (LSTM, por sus siglas en inglés). El modelo está diseñado con una capa de entrada, dos capas ocultas y una capa de salida. La capa de entrada recibe los datos preprocesados; las capas ocultas permiten reconocer patrones complejos en los datos (de 50 y 20 neuronas, respectivamente). Después de las capas ocultas, hay una capa de salida que clasifica cada entrada procesada en diez posibles gestos.

Inicialmente, se entrenó el modelo con tres millones de datos, de los cuales el 70% se destinó a la etapa de entrenamiento y el 30% restante se empleó para la validación de éste. El modelo LSTM alcanzó una exactitud de 0.99, lo que significa que 99% de las predicciones que el modelo hizo en los datos de validación fueron correctas. Posteriormente, se puso a prueba recibiendo datos sin procesar del Tap Strap 2 en tiempo real. De todas las veces que el modelo realizó la predicción para cada uno de los gestos definidos en el diccionario, acertó 99% de las veces.

Además, se hicieron ajustes para que las lecturas del Tap Strap 2 se adaptaran a cada usuario, teniendo en cuenta características como la forma de

su mano o la manera en que realiza los gestos. Este proceso, llamado corrección por factores, estabiliza la precisión del reconocimiento cuando el usuario que realiza un gesto no es la misma persona con la que se entrenó el modelo.



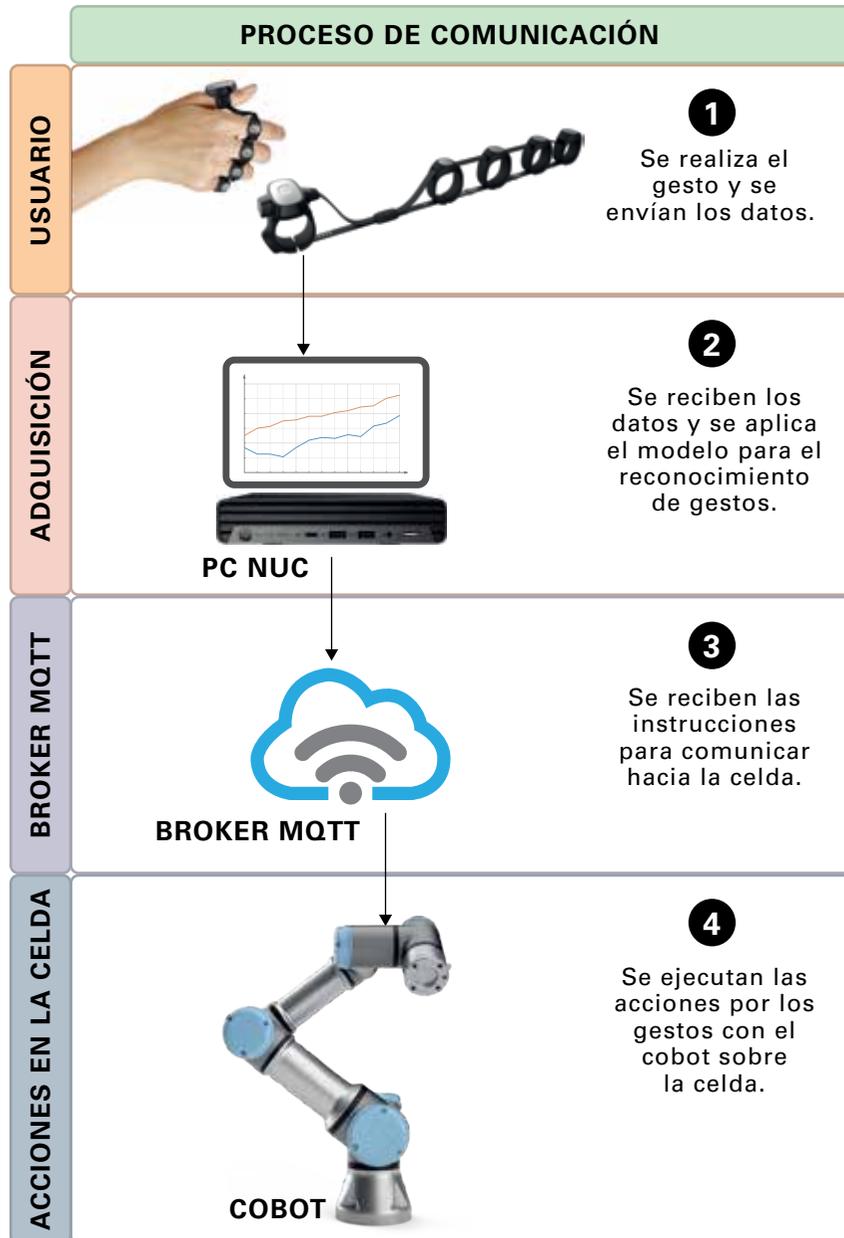


Figura 4. Diagrama de comunicación control gestual-celda de manufactura. Elaborado por los autores.

Comunicación

Una vez que el modelo de red neuronal fue implementado para clasificar gestos en tiempo real, el siguiente paso consistió en vincular cada gesto con una acción dentro de la celda de manufactura. Así, cada vez que se reconoce un gesto, se ejecuta una función específica que establece una conexión con el *broker*. Esta función envía una serie de mensajes en los tópicos correspondientes y el *broker* los reenvía a una máquina o dispositivo específico. Este cliente, al

recibir los mensajes, interpreta los valores y ejecuta la acción correspondiente; por ejemplo, mover un robot a una posición determinada. La Figura 4 describe gráficamente el proceso de la comunicación, sintetizando todos los aspectos abordados en el artículo.

Conclusiones

¿Imaginas controlar tecnología con gestos de la mano? El artículo presenta el desarrollo de un sis-

tema ciberfísico innovador basado en el reconocimiento de los gestos de la mano para transformar los sistemas de producción, lo que permite a personas con ciertas discapacidades físicas manejar y controlar la maquinaria de una celda de manufactura con sólo los movimientos de la mano.

Controlar tecnología con gestos podría transformar nuestra interacción con el mundo digital y físico. Este desarrollo, basado en un sistema que combina sensores portátiles (como el Tap Strap 2), inteligencia artificial y redes IoT, permitiría, por ejemplo, que personas con movilidad reducida operen dispositivos de forma autónoma, que trabajadores manipulen maquinaria peligrosa a distancia, o que médicos controlen equipos en quirófanos sin romper la esterilidad.

Además, al simplificar el uso de tecnología mediante gestos intuitivos, se promovería la inclusión digital de niños, adultos mayores o personas no expertas, reduciría la dependencia de dispositivos físicos (contribuyendo a la sostenibilidad) y crearía entornos más seguros y eficientes en industrias y hospitales. No se trata sólo de innovación: es un avance hacia una tecnología que se adapta a las necesida-

des humanas, priorizando accesibilidad, seguridad y bienestar colectivo.

Los autores agradecen al Laboratorio Nacional de Investigación en Tecnología Digital (LANITED) del Conahcyt y al programa Investigadores por México-Conahcyt por brindar la oportunidad de investigación a través del proyecto 730-2017. Esta investigación fue financiada por el Consejo Nacional de Humanidades, Ciencias y Tecnologías de México (Conahcyt) con el número de subvención F003-322609 (LANITED).

Ángel Iván García Moreno

Centro de Ingeniería y Desarrollo Industrial.
angel.garcia@cidesi.edu.mx

Ángel Gabriel Salinas Martínez

Centro de Ingeniería y Desarrollo Industrial.
angel.salinas@cidesi.edu.mx

Heber Isidro Morales Lugo

Centro de Ingeniería y Desarrollo Industrial.
heber.morales@cidesi.edu.mx

Lecturas recomendadas

Ceriani, A. (2020), “Los gestos y las máquinas en los nuevos comportamientos digitalizantes”, en M. Scarnatto y F. A. de Marziani (comps.), *Investigar en Cuerpo, Arte y Comunicación*. Disponible en: <https://www.teseopress.com/investigarencayc/chapter/los-gestos-y-las-maquinas-en-los-nuevos-comportamientos-digitalizantes/>, consultado el 30 de enero de 2025.

Mrazek, K., B. Holton, T. Klein, I. Khan, T. Ayele y T. Khan Mohd (2021), “The Tap Strap 2: Evaluating Performance of One-Handed Wearable Keyboard and Mouse”, *HCI International 2021-Late Breaking*

Papers: Multimodality, eXtended Reality, and Artificial Intelligence, pp. 82-95. Disponible en: https://doi.org/10.1007/978-3-030-90963-5_7, consultado el 30 de enero de 2025.

Trujillo-Romero, F. y G. García-Bautista (2021), “Reconocimiento de palabras de la Lengua de Señas Mexicana utilizando información RGB-D”, *Revista Electrónica de Computación, Informática, Biomédica y Electrónica*, 10(2):1-23. Disponible en: <https://www.redalyc.org/journal/5122/512269058001/>, consultado el 30 de enero de 2025.