



**Universidad  
Veracruzana**

**Facultad de Negocios y  
Tecnologías**

**INTELIGENCIA ARTIFICIAL**

**REPORTE DE  
PROYECTO**

**Profesor: Jesús Leonardo López Hernández**

**Estudiante: Imer Azareel Ortega Espejel**

**Fecha de entrega: Viernes 13 de Junio de 2025**

## TABLA DE CONTENIDO

TABLA DE ILUSTRACIONES .....	3
PLANTEAMIENTO DEL PROBLEMA .....	4
JUSTIFICACIÓN .....	5
ARQUITECTURA.....	6
Tecnologías implementadas .....	7
DESARROLLO .....	8
Elementos del prototipo desarrollado .....	9
IMPLEMENTACIÓN .....	11
PRUEBAS.....	13
CONCLUSIONES .....	16

## TABLA DE ILUSTRACIONES

Ilustración 1 Diagrama de tecnologías .....	6
Ilustración 2 Archivos del programa .....	12
Ilustración 3 Entrenamiento del modelo .....	13
Ilustración 4 tecla resltada .....	13
Ilustración 5 Letra agregada al teclado .....	14
Ilustración 6 Tecla borrar.....	14
Ilustración 7 Sugerencia de palabras .....	15
Ilustración 8 Teclado .....	15

## PLANTEAMIENTO DEL PROBLEMA

En la sociedad actual, la tecnología es una herramienta fundamental para la comunicación, el aprendizaje, el trabajo y el entretenimiento. Sin embargo, no todas las personas pueden acceder a ella de la misma manera. Las personas con discapacidad motriz, especialmente aquellas que tienen dificultades para mover las manos o los dedos, enfrentan grandes barreras para interactuar con computadoras, tabletas y teléfonos inteligentes. Los dispositivos tradicionales, como el teclado y el mouse, están diseñados para quienes pueden realizar movimientos precisos y coordinados, lo que deja fuera a quienes tienen movilidad reducida o nula en las extremidades superiores.

Aunque existen algunas soluciones tecnológicas para mejorar la accesibilidad, como teclados adaptados, dispositivos de seguimiento de cabeza o sistemas de control por voz, estas alternativas presentan varios inconvenientes. Muchas veces son costosas, requieren hardware especializado o no están disponibles en todos los lugares. Además, pueden ser difíciles de configurar y usar, especialmente para personas que no tienen experiencia con tecnología. En el caso de los sistemas de control por voz, por ejemplo, pueden fallar en ambientes ruidosos o no reconocer correctamente el habla de todos los usuarios.

En este contexto, el seguimiento de la mano, y en particular del dedo índice, surge como una opción innovadora y accesible. Gracias a los avances en visión por computadora, es posible utilizar una cámara web común para detectar el movimiento del dedo índice y traducirlo en acciones sobre un teclado virtual en la pantalla. De esta manera, una persona puede escribir texto, seleccionar opciones y comunicarse sin necesidad de usar un teclado físico. Sin embargo, el desarrollo de este tipo de sistemas no está exento de retos. Es necesario lograr que el seguimiento sea preciso y rápido, que el sistema sea fácil de usar y que funcione en diferentes condiciones de iluminación y con distintos usuarios. Además, es importante que la herramienta sea económica y que no requiera equipos difíciles de conseguir.

Por todo lo anterior, el problema principal que se busca resolver es la falta de una solución accesible, económica y fácil de usar que permita a las personas con discapacidad motriz escribir y comunicarse usando solo el movimiento de su dedo índice. Superar este reto es fundamental para reducir la brecha digital y mejorar la calidad de vida de quienes enfrentan barreras físicas para interactuar con la tecnología.

## JUSTIFICACIÓN

La creación de una herramienta que permita a las personas con discapacidad motriz escribir y comunicarse usando únicamente el movimiento de su dedo índice está plenamente justificada por varias razones de peso. En primer lugar, la tecnología debe ser una aliada para la inclusión, no una barrera. Si bien existen soluciones comerciales para la accesibilidad, muchas de ellas son costosas, requieren hardware especial o no se adaptan a las necesidades de todos los usuarios. Un sistema basado en el seguimiento del dedo índice, utilizando una cámara web común y software de visión por computadora, representa una alternativa mucho más accesible y adaptable.

Este tipo de herramienta puede transformar la vida de las personas con discapacidad motriz. Al permitirles escribir texto y seleccionar palabras en una pantalla usando solo el dedo índice, se les brinda una mayor independencia y autonomía. Ya no dependerán de la ayuda constante de otras personas para comunicarse o realizar tareas básicas en la computadora. Además, podrán participar en actividades educativas, laborales y sociales que antes les estaban vedadas, lo que contribuye a su integración y bienestar.

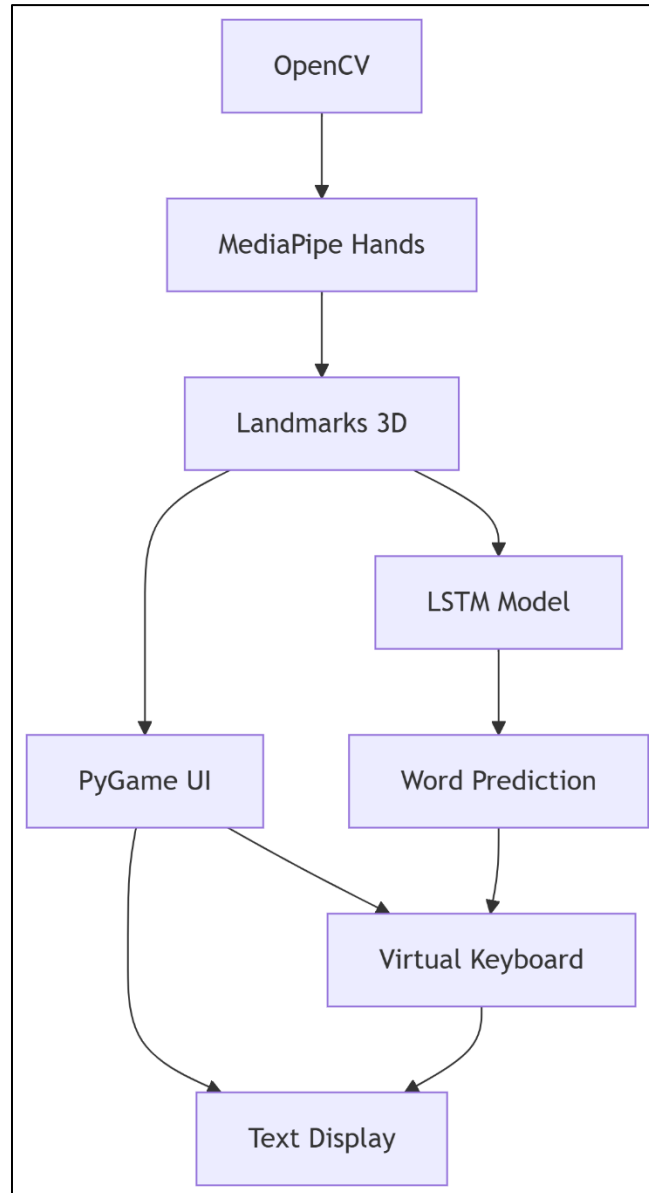
Otra razón importante para desarrollar este sistema es que puede adaptarse a diferentes contextos y necesidades. Al estar basado en software y hardware común, puede instalarse en computadoras personales, laptops o incluso tabletas, sin necesidad de comprar equipos costosos o difíciles de conseguir. Además, el sistema puede personalizarse para ajustarse a las capacidades y preferencias de cada usuario, por ejemplo, cambiando el tamaño de las teclas, la sensibilidad del seguimiento o el tipo de sugerencias de palabras.

El impacto social de una herramienta así es significativo. Muchas personas con discapacidad motriz experimentan aislamiento social y dificultades para acceder a la educación o al empleo debido a las barreras tecnológicas.

Finalmente, el desarrollo de este tipo de tecnología es un compromiso ético y social. Promover la igualdad de oportunidades y la inclusión de las personas con discapacidad es responsabilidad de todos. La tecnología debe ser una herramienta para derribar barreras, no para crearlas. Al invertir en soluciones accesibles y adaptables, se contribuye a la dignidad y los derechos de las personas con discapacidad, permitiéndoles expresarse, aprender y participar plenamente en la vida digital.

## ARQUITECTURA

El sistema sigue un flujo ordenado desde que el usuario mueve su dedo hasta que se genera texto en pantalla. Aquí está el desglose visual simplificado:



*Ilustración 1 Diagrama de tecnologías*

## Tecnologías implementadas

### Módulo de Visión por Computadora

- Lenguaje/Framework: Python + OpenCV + MediaPipe
- Funcionalidad:
  - Captura de video en tiempo real mediante una cámara web estándar.
  - Detección y seguimiento del dedo índice usando el modelo MediaPipe Hands.
  - Procesamiento de *landmarks* (21 puntos clave por mano) para determinar la posición (x, y) del dedo en la pantalla.
  - Salida: Coordenadas normalizadas del dedo índice para interactuar con la UI.

### Módulo de Lógica de Teclado

- Lenguaje/Framework: Python + PyTorch
- Funcionalidad:
  - Selección de Teclas: Mapeo de coordenadas del dedo a teclas virtuales.
  - Predicción de Palabras: Modelo LSTM entrenado con vocabulario en español para autocompletado.
  - Gestión de Estado: Control de mayúsculas, borrado y selección de sugerencias.
  - Salida: Texto generado y sugerencias para el usuario.

### 3. Interfaz de Usuario (UI)

- Lenguaje/Framework: PyGame
- Componentes
  - Teclado Virtual: Distribución QWERTY adaptada con teclas grandes (60x60 px).
  - Área de Texto: Muestra el texto ingresado en tiempo real.
  - Sugerencias: Barra inferior con 3 palabras predichas (resaltadas al detectar gestos).
- Feedback Visual:
  - Resaltado de teclas al seleccionarlas.
  - Cursor intermitente para orientación.

## DESARROLLO

El desarrollo del prototipo se llevó a cabo en varias etapas, siempre pensando en que fuera fácil de usar y realmente útil para personas con discapacidad motriz. El objetivo principal fue crear una herramienta que permitiera escribir en una computadora usando únicamente el movimiento del dedo índice, sin necesidad de un teclado físico ni de equipos costosos.

- Investigación y planeación

En un primer momento, se investigaron las dificultades que enfrentan las personas con discapacidad motriz para escribir o comunicarse usando una computadora. Se revisaron diferentes soluciones existentes, como teclados especiales, sistemas de control por voz y otras tecnologías de asistencia. Sin embargo, se notó que muchas de estas opciones son caras, difíciles de conseguir o no se adaptan bien a todos los usuarios. Por eso, se decidió buscar una alternativa más accesible y sencilla, usando el seguimiento del dedo índice a través de una cámara web común.

- Diseño del teclado virtual

Después de definir el enfoque, se diseñó un teclado virtual que aparece en la pantalla de la computadora. Este teclado tiene todas las letras del abecedario, los números y algunas teclas especiales como “borrar”, “espacio” y “mayúsculas”. El diseño se pensó para que las teclas fueran lo suficientemente grandes y separadas, facilitando que el usuario pueda seleccionarlas con el dedo índice sin dificultad. Además, se agregó una zona en la parte superior para mostrar el texto que el usuario va escribiendo y otra zona para mostrar sugerencias de palabras, ayudando a escribir más rápido y con menos esfuerzo.

- Implementación del seguimiento del dedo índice

En esta fase, se utilizó una cámara web común y un programa de visión por computadora para detectar la mano del usuario y, en especial, la posición del dedo índice. El sistema reconoce en tiempo real dónde está el dedo índice y traduce ese movimiento a una posición en la pantalla. Cuando el dedo índice se coloca sobre una tecla del teclado virtual, el sistema lo interpreta como una selección y escribe la letra correspondiente en la zona de texto. Para borrar una letra, el usuario solo necesita colocar el dedo índice sobre la tecla de borrar.

Además, se pensó en una función especial para borrar palabras completas. Si el usuario junta el dedo índice y el pulgar durante cuatro segundos sobre la tecla de borrar, el sistema borra la palabra anterior de una sola vez. Esto facilita la corrección de errores y hace que el proceso de escritura sea más rápido y cómodo.



- Sistema de sugerencias de palabras

Para hacer la escritura aún más eficiente, se desarrolló un sistema de sugerencias de palabras. Se entrenó una red neuronal con un archivo de palabras en español, de modo que el sistema pueda predecir y sugerir palabras completas a medida que el usuario va escribiendo letras. Estas sugerencias aparecen en la zona de sugerencias y el usuario puede seleccionarlás colocando el dedo índice sobre la palabra sugerida. Así, se reduce el número de movimientos necesarios para escribir palabras largas o comunes.

- Integración y pruebas

Una vez que todos los componentes estuvieron listos, se integraron en un solo programa. Se realizaron varias pruebas para asegurarse de que el sistema respondiera correctamente al movimiento del dedo índice y a la acción de juntar el índice y el pulgar para borrar palabras completas. También se ajustaron detalles como el tamaño de las teclas, la sensibilidad del seguimiento y el tiempo necesario para activar el borrado de palabras, buscando siempre que el sistema fuera cómodo y fácil de usar para cualquier persona.

## Elementos del prototipo desarrollado

- Teclado virtual: Un teclado que aparece en la pantalla y permite escribir letras, números y símbolos usando solo el dedo índice.
- Zona de texto: Un área donde se muestra en tiempo real lo que el usuario va escribiendo.
- Zona de sugerencias: Un espacio donde aparecen palabras completas que el usuario puede seleccionar para escribir más rápido.
- Seguimiento del dedo índice: El sistema detecta la posición del dedo índice y permite seleccionar teclas sin necesidad de un teclado físico.
- Función de borrado inteligente: Si el usuario junta el dedo índice y el pulgar durante cuatro segundos sobre la tecla de borrar, se borra la palabra completa anterior.
- Red neuronal entrenada: El sistema sugiere palabras basándose en las letras que el usuario va escribiendo, haciendo la escritura más eficiente y rápida.

El prototipo desarrollado es una herramienta accesible y fácil de usar, pensada para personas con discapacidad motriz. Permite escribir y comunicarse usando solo el movimiento del dedo índice, detectado por una cámara web común. Gracias a su diseño sencillo y a sus funciones inteligentes, el sistema ayuda a mejorar la autonomía y la inclusión digital de los usuarios, permitiéndoles participar activamente en la sociedad digital sin depender de equipos costosos o difíciles de conseguir.

## IMPLEMENTACIÓN

El código del proyecto está dividido en varios archivos y utiliza librerías como MediaPipe, OpenCV, PyTorch y Pygame. Cada archivo y función cumple un papel importante para que el sistema funcione correctamente.

Para el seguimiento del dedo índice, se usa MediaPipe junto con OpenCV. En el código, una función procesa cada imagen de la cámara y obtiene los puntos clave de la mano. Por ejemplo, la función `is_index_thumb_together()` calcula la distancia entre el dedo índice y el pulgar usando los landmarks que da MediaPipe. Así, el sistema puede detectar gestos especiales, como el de borrar una palabra completa.

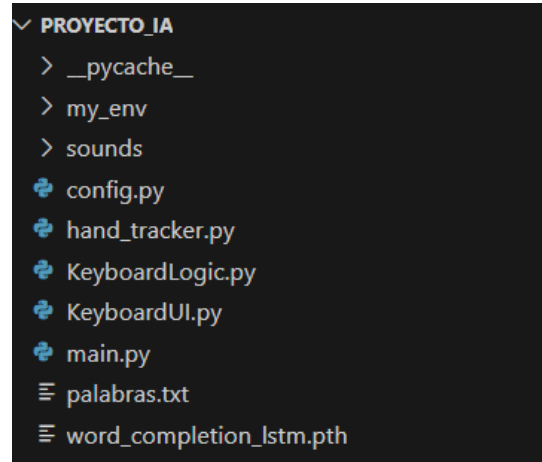
En el archivo de lógica del teclado, la función `update_selection(pos)` revisa si la posición del dedo índice coincide con alguna tecla del teclado virtual. Si es así, la función `select_key(key, borrar_palabra_completa=False)` agrega la letra al texto o borra un carácter. Si el parámetro `borrar_palabra_completa` es verdadero, borra toda la palabra anterior. Estas funciones también manejan teclas como "ESPACIO" y "MAYUS".

Para las sugerencias de palabras, el código utiliza **PyTorch** para crear, entrenar y cargar un modelo de red neuronal tipo LSTM. El entrenamiento del modelo se realiza al inicio del programa, usando un archivo de palabras en español como base de datos. El proceso de entrenamiento consiste en tomar cada palabra y generar ejemplos de prefijos y palabras completas, que luego se usan para enseñar al modelo a predecir cómo completar una palabra a partir de sus primeras letras. Durante el entrenamiento, PyTorch gestiona la creación de la red neuronal, la función de pérdida y el optimizador. En la consola, el sistema muestra mensajes que indican el tamaño del vocabulario, la cantidad de ejemplos generados y el avance de cada época de entrenamiento. Una vez entrenado, el modelo se guarda en un archivo para que no sea necesario volver a entrenarlo cada vez que se ejecuta el programa.

La función `_predict_words(prefix, n=3)` toma las letras que el usuario ha escrito y le pide al modelo que sugiera palabras completas que empiecen igual. La función `get_suggestions(n=3)` muestra estas sugerencias en la pantalla.

La interfaz gráfica está hecha con Pygame. Funciones como `_draw_keyboard()`, `_draw_text_area()` y `_draw_suggestions()` dibujan el teclado virtual, el área de texto y las sugerencias en la pantalla. Cada vez que el usuario mueve el dedo o selecciona una tecla, la interfaz se actualiza para mostrar el cambio.

El archivo principal del programa, normalmente llamado `main.py`, coordina todo el sistema. Usa un bucle que recibe la imagen de la cámara, llama a las funciones de MediaPipe para detectar la mano, actualiza la lógica del teclado y redibuja la interfaz con Pygame. Así, el usuario puede escribir, borrar y seleccionar sugerencias usando solo el dedo índice y gestos sencillos.



*Ilustración 2 Archivos del programa*

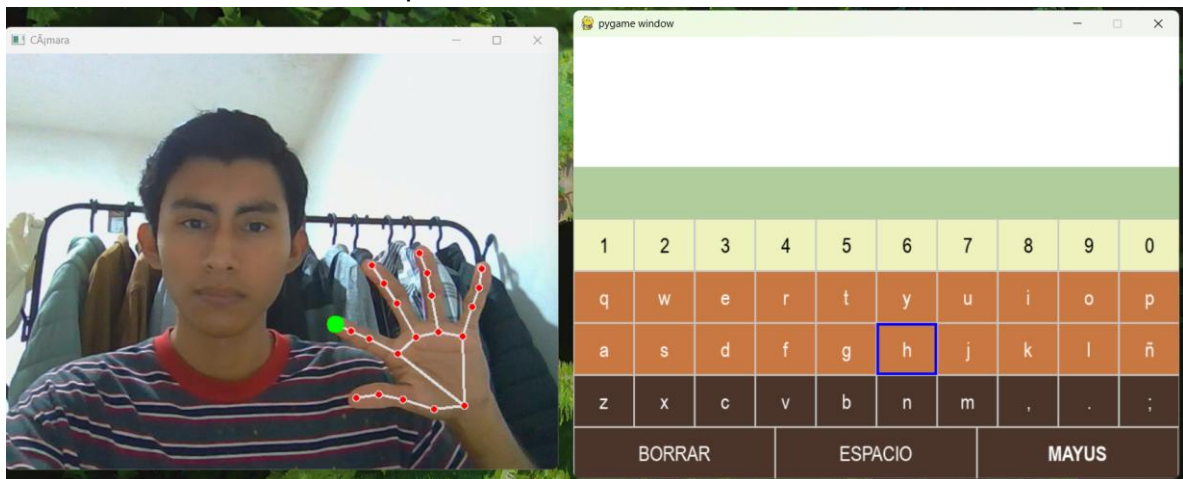
## PRUEBAS

Al ejecutar el archivo principal (main.py), el sistema inicia el entrenamiento de una red neuronal para el autocompletado de palabras. Durante este proceso, en la consola se muestran mensajes que indican el avance, como el tamaño del vocabulario, la cantidad de pares prefijo-palabra y el progreso de las épocas de entrenamiento. Una vez finalizado el entrenamiento, el sistema carga la interfaz gráfica del teclado virtual y comienza la captura de video para el seguimiento del dedo índice.

```
(my_env) PS C:\Users\iaoe0\Documents\Proyecto_ia> python main.py
pygame 2.6.1 (SDL 2.28.4, Python 3.11.0)
Hello from the pygame community. https://www.pygame.org/contribute.html
Cargando palabras...
Tamaño del vocabulario de caracteres: 32
Número de palabras usadas para entrenar: 9952
Entrenando modelo LSTM...
Época 1/10
Época 2/10
Época 3/10
Época 4/10
Época 5/10
Época 6/10
Época 7/10
```

*Ilustración 3 Entrenamiento del modelo*

- Al mover el dedo índice frente a la cámara, el sistema detecta la posición y resalta la tecla virtual correspondiente en el teclado.



*Ilustración 4 tecla resaltada*

- Al mantener el dedo índice sobre una tecla durante el tiempo necesario, el carácter de esa tecla se agrega al área de texto en la parte superior de la pantalla.

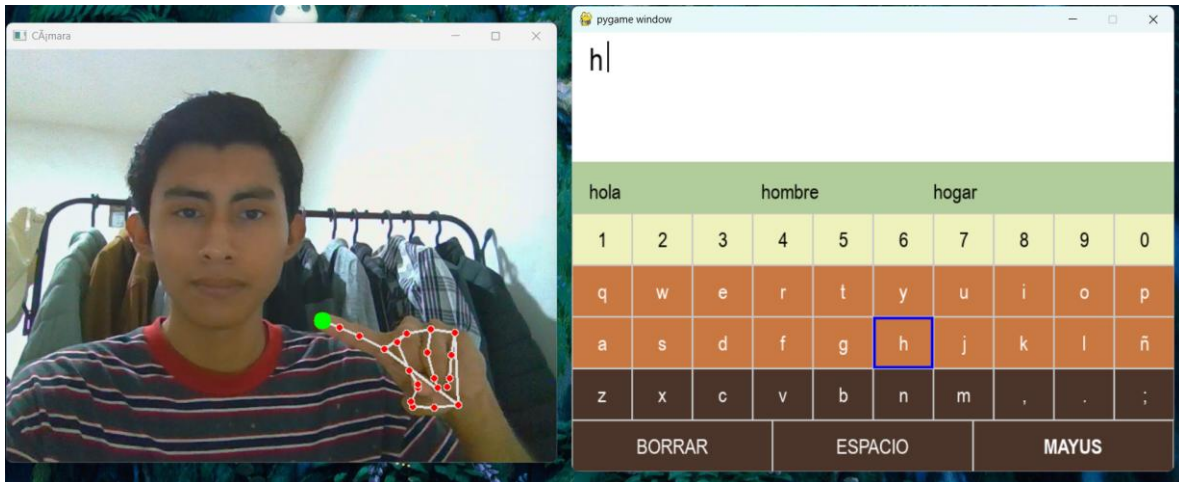


Ilustración 5 Letra agregada al teclado

- Al posicionar el dedo índice sobre la tecla de borrar, el sistema elimina el último carácter del texto. Cada repetición de esta acción borra un carácter adicional.

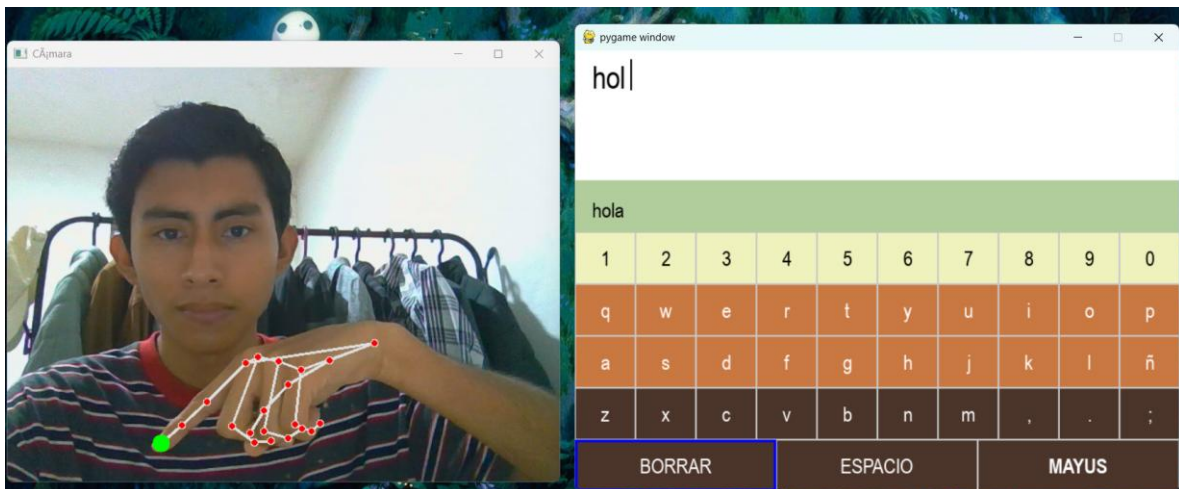


Ilustración 6 Tecla borrar

- Al escribir las primeras letras de una palabra, el sistema muestra sugerencias de palabras completas en la zona de sugerencias. Al posicionar el dedo índice sobre una sugerencia, la palabra sugerida se completa automáticamente en el área de texto.

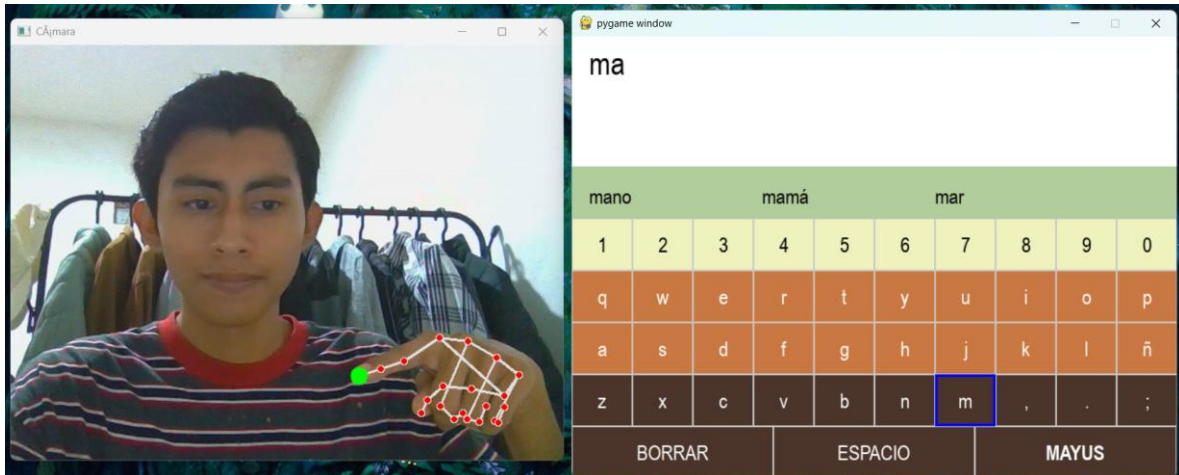


Ilustración 7 Sugerencia de palabras

- Estos casos de prueba validan que el sistema realiza el entrenamiento inicial del modelo, reconoce el movimiento del dedo índice, permite la selección de caracteres, el uso de sugerencias y el borrado de texto, todo mediante gestos simples frente a la cámara.

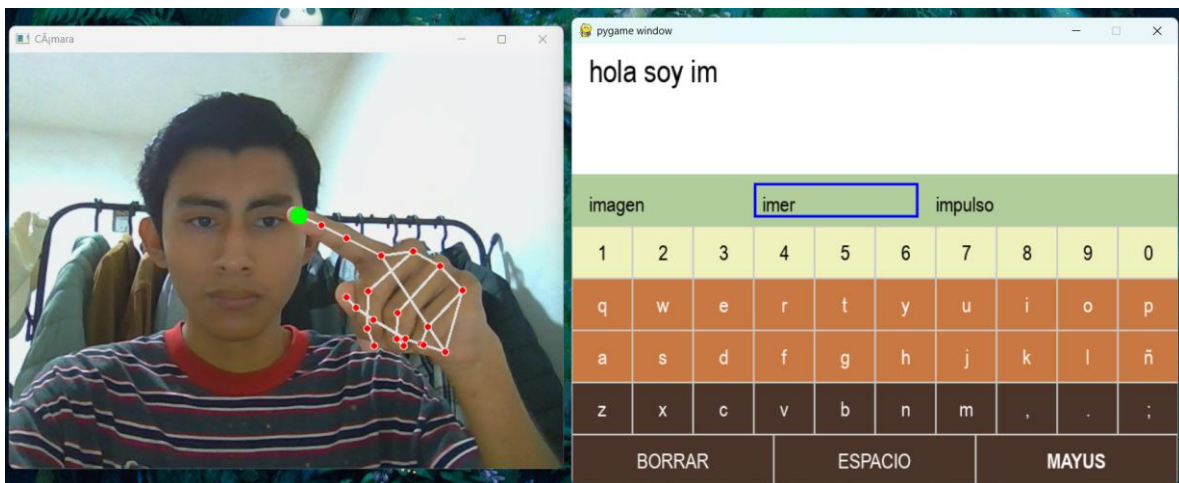


Ilustración 8 Teclado



## CONCLUSIONES

El teclado virtual basado en el seguimiento del dedo índice ha demostrado ser una herramienta útil para facilitar la escritura y la comunicación en personas con discapacidad motriz. El sistema permite seleccionar caracteres, borrar texto y utilizar sugerencias de palabras mediante gestos sencillos frente a la cámara, sin requerir dispositivos especiales. Sin embargo, el prototipo aún presenta algunas limitaciones y existen varias áreas en las que se puede mejorar.

Entre las limitaciones más notables se encuentra el modelo de autocompletado, que todavía no sugiere una gran variedad de palabras y requiere un entrenamiento más profundo y con un vocabulario más amplio para ofrecer mejores resultados. Además, la detección de gestos está limitada principalmente al movimiento del dedo índice.

Para futuras mejoras, se propone implementar nuevos gestos que permitan realizar más acciones, como activar el modo de mayúsculas, insertar espacios o borrar palabras completas de manera más eficiente. También sería beneficioso optimizar el modelo de autocompletado, entrenándolo con un dataset de palabras más grande y variado, para que las sugerencias sean más útiles y adaptadas al contexto del usuario. Finalmente, se podrían explorar opciones de personalización en la interfaz, permitiendo ajustar el tamaño de las teclas o los colores, y agregar un sistema de calibración para adaptar la sensibilidad del seguimiento a cada usuario.

En conclusión, aunque el prototipo cumple con su objetivo principal, todavía existen retos y oportunidades para hacerlo más completo, preciso y adaptable a las necesidades de los usuarios.