



**Universidad Veracruzana
Región Orizaba - Córdoba
Campus de Ixtaczoquitlán
Facultad de Negocios y Tecnologías**

Programa educativo:
Tecnologías de Información en las Organizaciones

Alumno:
Jasub Torres Ramírez

Experiencia educativa:
Inteligencia Artificial

Profesor:
Jesús Leonardo López Hernández

Actividad:
Proyecto Final

Fecha de entrega:
13 de junio de 2025

Tabla de contenido

1.	Introducción	3
2.	Arquitectura del Sistema.....	4
2.1	Diagrama de Componentes	4
3.	Desarrollo y Prototipo	5
3.1	Herramientas y Tecnologías Utilizadas.....	5
3.2	Estructura del Prototipo	5
3.3	Interacción sin contacto y retroalimentación.....	6
3.4	Personalización y Adaptabilidad.....	6
4.	Implementación (Funciones Clave).....	7
4.1	Inicialización y Configuración.....	7
4.2	Detección de Mano e Índice	7
4.3	Limpieza del Texto con Gesto de Mano Abierta	8
4.4	Interfaz del Teclado Virtual.....	8
4.5	Selección de Teclas.....	8
4.6	Predicción de Palabras	9
4.7	Interfaz Principal	9
5.	Pruebas Realizadas	10
5.1	Prueba de Detección de Mano y Punta del Dedo Índice	10
5.2	Prueba de Selección de Teclas	11
5.3	Prueba del Gesto de Palma Abierta (Reset de Texto)	12
5.4	Prueba del Sistema de Sugerencias	13
5.5	Prueba de Distribución Visual y Ergonomía	14
6.	Conclusiones	15

1. Introducción

En la actualidad, la tecnología desempeña un papel fundamental en la comunicación, el aprendizaje, el trabajo y el acceso a la información. Sin embargo, muchas de estas tecnologías no están diseñadas considerando las necesidades de las personas con discapacidades físicas o motoras. Una de las barreras más comunes que enfrentan las personas con discapacidad motriz parcial o severa es la interacción con interfaces tradicionales como el teclado y el ratón. Estas herramientas requieren de habilidades motoras finas y coordinación que no todas las personas poseen, lo que limita su capacidad de escribir, navegar por internet o utilizar aplicaciones cotidianas.

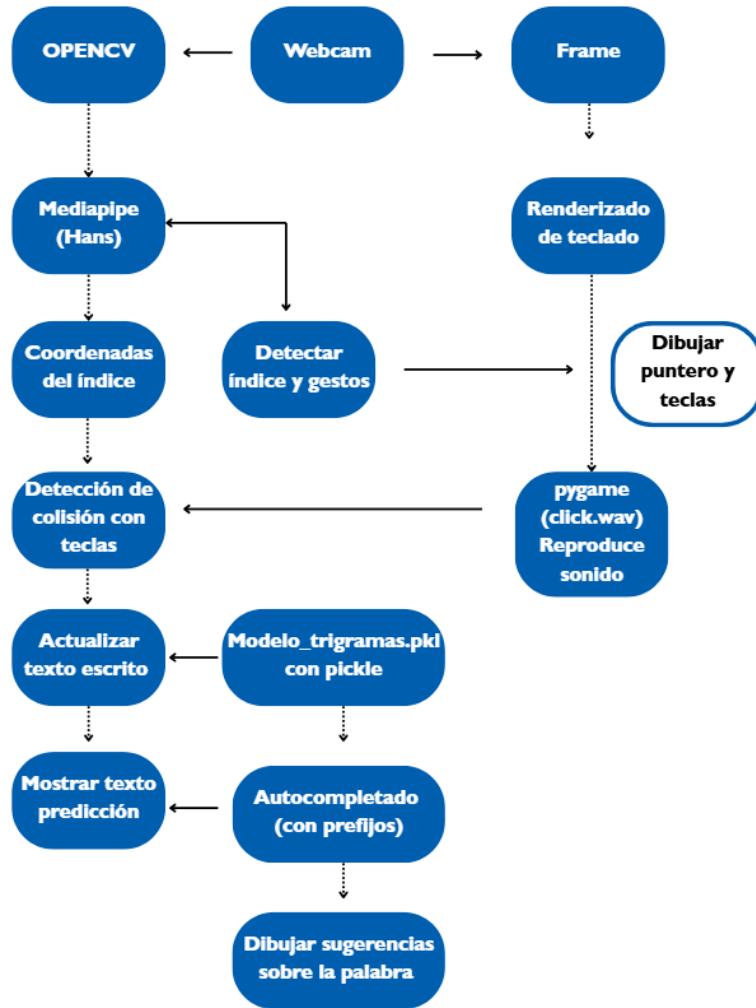
A pesar de que existen tecnologías de asistencia en el mercado, como teclados especiales, sistemas de control por voz o seguimiento ocular (eye tracking), muchas de estas soluciones son costosas, complejas de configurar o poco accesibles para la mayoría de las personas. Además, algunas de ellas dependen de hardware especializado, lo cual las hace inaccesibles en contextos de bajos recursos o en comunidades marginadas. Ante esta situación, surge la necesidad de desarrollar alternativas más inclusivas, asequibles y efectivas que permitan a las personas con discapacidad comunicarse e interactuar con dispositivos electrónicos de manera autónoma.

Este proyecto surge con la finalidad de ofrecer una solución innovadora y accesible basada en Inteligencia Artificial y visión por computadora. En lugar de depender de dispositivos costosos o especializados, se utiliza una cámara web estándar para detectar y seguir el movimiento de la mano del usuario. A través de una interfaz visual amigable que emula un teclado, el sistema permite que el usuario escriba texto moviendo su mano frente a la cámara. La selección de caracteres se realiza cuando el dedo índice permanece sobre una tecla virtual durante un determinado tiempo, lo que elimina la necesidad de presionar botones físicos. Además, se incorporó un módulo de predicción de palabras basado en trigramas para facilitar la escritura y reducir el número de interacciones necesarias.

Este enfoque no solo demuestra la aplicabilidad de la Inteligencia Artificial en contextos reales y sociales, sino que también refuerza el compromiso con la inclusión digital. Al permitir que una persona con movilidad reducida pueda redactar mensajes, buscar información o comunicarse por medios digitales sin dispositivos físicos adicionales, se está abriendo una puerta hacia una participación más plena en la vida académica, laboral y personal. En este sentido, el proyecto busca no solo ser una solución tecnológica, sino también una herramienta de empoderamiento y autonomía.

2. Arquitectura del Sistema

2.1 Diagrama de Componentes



Tecnologías implementadas:

- OpenCV: Captura de video, detección de interacción con teclas y dibujo del teclado.
- MediaPipe (Hands): Detección de landmarks de la mano y gestos como la palma abierta.
- Pygame: Sonido de retroalimentación al presionar una tecla.
- Pickle: Carga del modelo de lenguaje basado en trigramas para autocorreo.
- Python: Lenguaje de programación principal.

3. Desarrollo y Prototipo

El desarrollo del prototipo se centró en construir una interfaz funcional que permitiera a una persona con movilidad reducida escribir texto utilizando solo el movimiento de una mano frente a una cámara. A diferencia del planteamiento original basado en el seguimiento ocular y detección de parpadeos, este prototipo se rediseñó con un enfoque en **visión por computadora basada en las manos**, haciendo uso de una cámara convencional y bibliotecas especializadas de inteligencia artificial.

3.1 Herramientas y Tecnologías Utilizadas

El proyecto se desarrolló en Python, utilizando las siguientes librerías y recursos principales:

- **OpenCV**: para el procesamiento de video en tiempo real y renderizado de la interfaz visual.
- **MediaPipe Hands**: un modelo preentrenado de Google que permite detectar la mano humana y obtener las coordenadas 3D de puntos clave como dedos, palma y articulaciones.
- **Pygame**: para la reproducción de sonidos que ofrecen retroalimentación al usuario al seleccionar una tecla.
- **Modelo de Trigramas (pickle)**: un modelo básico de predicción de palabras implementado a partir de secuencias de texto, cargado desde un archivo modelo_trigramas.pkl entrenado previamente.
- **Webcam estándar**: para capturar en tiempo real la posición de la mano.

3.2 Estructura del Prototipo

El sistema está compuesto por tres componentes principales:

1. Sistema de detección de mano y punto índice.

Utilizando MediaPipe Hands, se detecta la presencia de una mano en el marco de video. A partir de los puntos clave proporcionados por la librería, se extrae la posición del dedo índice (punto 8), que se utiliza como cursor sobre la pantalla. Esta técnica permite una interacción natural y sin contacto con la interfaz.

2. Interfaz gráfica del teclado virtual.

Se diseñó un teclado en pantalla dividido en varias filas con letras y funciones básicas como "SPACE", "DEL", y números del 0 al 3. Cada tecla tiene un área rectangular que es dibujada

y renderizada en tiempo real. El sistema detecta si el dedo índice se encuentra sobre una tecla durante más de un segundo, lo que se interpreta como una selección o “clic virtual”. Al presionar una tecla:

- Se reproduce un sonido de confirmación.
- Se añade el carácter correspondiente al texto escrito.
- Si la tecla es "SPACE", se inserta un espacio.
- Si la tecla es "DEL", se borra el último carácter.
- **Módulo de predicción de palabras**

Al escribir, el sistema genera sugerencias de palabras basadas en un modelo de trigramas. Estas predicciones se muestran visualmente como opciones que el usuario puede seleccionar también con la mano. Esto permite acelerar el proceso de escritura, especialmente para personas con movilidad limitada, al reducir la necesidad de escribir letra por letra.

3.3 Interacción sin contacto y retroalimentación

Una característica importante del prototipo es que no requiere de contacto físico ni de hardware especializado. La selección se basa únicamente en la permanencia del dedo índice sobre una tecla. Además, se implementó retroalimentación sonora mediante pygame, de forma que el usuario pueda saber cuándo se ha activado una tecla correctamente. Esto mejora significativamente la experiencia de uso, especialmente en personas con discapacidad visual parcial.

3.4 Personalización y Adaptabilidad

Durante el desarrollo, se hicieron ajustes visuales al teclado para mejorar la legibilidad y la alineación de los elementos. Por ejemplo, se redujo el tamaño de algunas teclas como "SPACE", se ajustó el espaciado entre columnas y se rediseñó la disposición de las filas para evitar que se vieran en forma de "escalera". Estos detalles son cruciales en interfaces accesibles, ya que deben ser simples, intuitivas y cómodas para la navegación visual.

4. Implementación (Funciones Clave)

La implementación del prototipo se realizó en Python, integrando diferentes módulos de inteligencia artificial y visión por computadora para lograr una interacción efectiva entre la mano del usuario y una interfaz gráfica de teclado virtual. A continuación, se describen las funciones y componentes principales del sistema.

4.1 Inicialización y Configuración

Al inicio del programa se cargan los módulos necesarios:

- cv2 para captura y procesamiento de video.
- mediapipe para el modelo de detección de manos.
- pygame para el sonido de clics al presionar teclas.
- pickle para cargar el modelo de predicción de trigramas previamente entrenado.

Además, se inicializan parámetros globales como el diseño del teclado, el tamaño de las teclas, y las variables para almacenar el texto escrito y las sugerencias.

4.2 Detección de Mano e Índice

Se utiliza MediaPipe Hands para detectar la mano en tiempo real. Luego se extrae la posición del dedo índice mediante la función:

```
def detectar_punta_index(hand_landmarks, width, height):  
    x = int(hand_landmarks.landmark[8].x * width)  
    y = int(hand_landmarks.landmark[8].y * height)  
    return x, y
```

Este punto se utiliza como un "puntero virtual" que navega sobre el teclado dibujado en pantalla.

4.3 Limpieza del Texto con Gesto de Mano Abierta

Se implementó una función que detecta si los dedos están abiertos, y si el gesto se detecta, se limpia el texto escrito:

```
def detectar_gesto_palma_abierta(landmarks):
    dedos_abiertos = [...]
    return all(dedos_abiertos) and (pulgar_izq or pulgar_der)
```

Este gesto se interpreta como una orden de “reiniciar”, útil para eliminar todo el contenido si el usuario desea comenzar de nuevo.

4.4 Interfaz del Teclado Virtual

El teclado se genera dinámicamente a partir de una lista de filas, cada una con sus respectivas letras o comandos. Se calcula su posición en pantalla y se dibuja con:

```
cv2.rectangle(...) # Para dibujar teclas
cv2.putText(...)    # Para mostrar el texto de la tecla
```

Además, se ajustó visualmente el tamaño y la alineación de teclas especiales como "SPACE" y "DEL" para que no se sobrepongan ni rompan la armonía visual.

4.5 Selección de Teclas

Cuando el dedo índice se posiciona sobre una tecla durante más de un segundo, se considera que la tecla ha sido “presionada”. Esto se controla con la variable `tiempo_inicio` y la comparación de tiempo transcurrido. Las acciones posibles incluyen:

- Agregar un carácter (`texto_escrito += tecla`)
- Insertar un espacio (`texto_escrito += " "`)
- Eliminar el último carácter (`texto_escrito = texto_escrito[:-1]`)
- Reproducir un sonido de confirmación (`click_sound.play()`)

Este mecanismo permite simular una experiencia de tecleo sin necesidad de botones físicos.

4.6 Predicción de Palabras

Una de las funciones clave del prototipo es la predicción de palabras para facilitar la escritura. La función predecir_palabra analiza el texto escrito y sugiere posibles continuaciones usando un modelo de trigramas cargado desde archivo:

```
def predecir_palabra(texto, modelo):
    palabras = texto.strip().split()
    ...
    return sorted(opciones)[:3]
```

Las sugerencias se dibujan en pantalla, y el usuario puede seleccionarlas con el mismo método de permanencia del dedo sobre ellas.

4.7 Interfaz Principal

La interfaz del sistema combina todos estos elementos en un bucle while True que:

1. Captura el video desde la cámara.
2. Procesa el marco para detectar manos.
3. Calcula las posiciones de las teclas y del dedo.
4. Dibuja el teclado y el texto escrito.
5. Detecta gestos y selecciona teclas o sugerencias.
6. Muestra todo en una ventana de OpenCV (cv2.imshow).

El programa finaliza cuando se presiona la tecla ESC.

5. Pruebas Realizadas

Para comprobar la funcionalidad, precisión y usabilidad del prototipo, se realizaron una serie de pruebas enfocadas en los componentes clave del sistema: la detección de mano, la interacción con las teclas, la selección mediante permanencia y el sistema de autocompletado de palabras.

5.1 Prueba de Detección de Mano y Punta del Dedo Índice

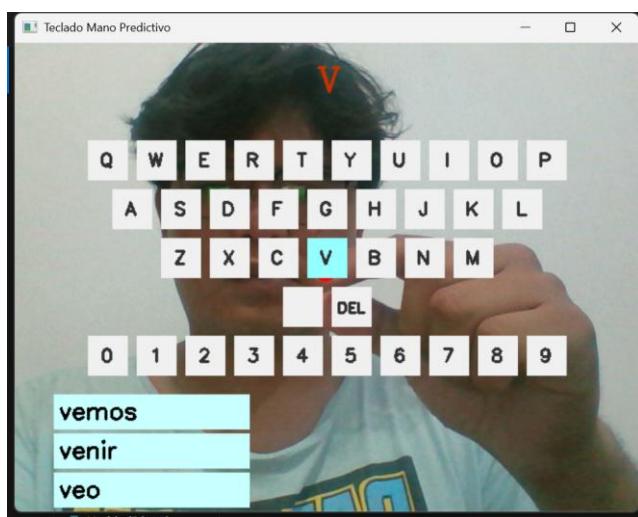
Objetivo: Verificar que el sistema detecte correctamente la posición del dedo índice en diferentes condiciones de iluminación, ángulos y distancias.

Procedimiento:

- Se encendió la cámara y se realizaron movimientos suaves del dedo índice frente a la interfaz.
- Se observaron las coordenadas de la punta del dedo y su representación visual (círculo rojo).
- Se probó en ambientes iluminados naturalmente y con luz artificial tenue.

Resultado esperado: El sistema detecta y sigue de forma estable la punta del dedo índice sin saltos o errores significativos.

Resultado obtenido: Cumplido. La detección se mantuvo precisa en un rango amplio de condiciones.



5.2 Prueba de Selección de Teclas

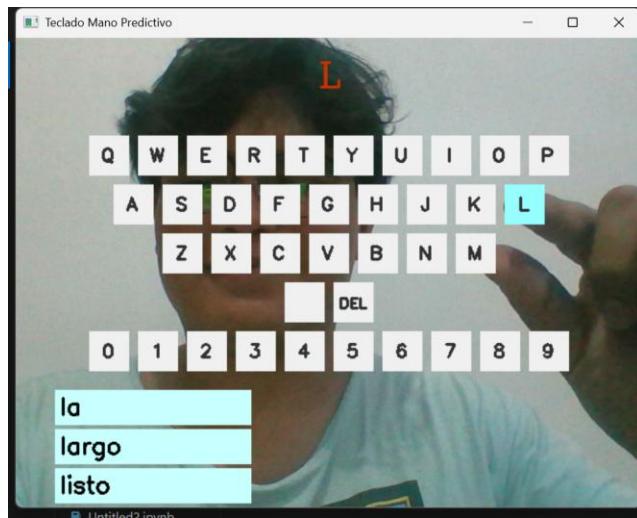
Objetivo: Validar que el usuario puede “presionar” una tecla al mantener el dedo sobre ella por más de un segundo.

Procedimiento:

- Se colocó el dedo sobre teclas específicas (como “A”, “B”, “DEL”, “SPACE”) por más de 1 segundo.
- Se observó si se agregaba el carácter deseado o si se eliminaba correctamente el último.

Resultado esperado: La tecla se selecciona una sola vez por permanencia, se escucha el sonido de confirmación y se actualiza el texto.

Resultado obtenido: Cumplido. El tiempo de retención de 1 segundo resultó intuitivo y efectivo para evitar selecciones accidentales.



5.3 Prueba del Gesto de Palma Abierta (Reset de Texto)

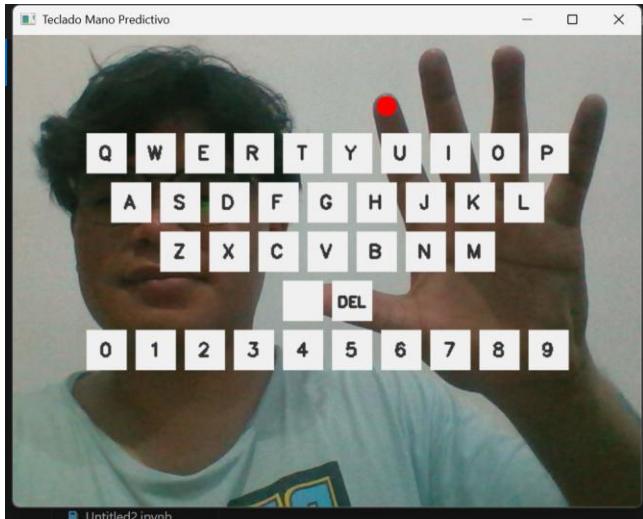
Objetivo: Verificar que al abrir completamente la mano frente a la cámara se borre todo el texto escrito, como mecanismo rápido de reinicio.

Procedimiento:

- Despues de escribir algunas palabras, se mostró la palma abierta frente a la cámara durante 1 segundo.
- Se observó si el texto desaparecía y las sugerencias se reiniciaban.

Resultado esperado: Se borra el texto completo.

Resultado obtenido: Cumplido. El gesto es fácil de realizar y funciona como mecanismo útil para borrar sin complicaciones.



5.4 Prueba del Sistema de Sugerencias

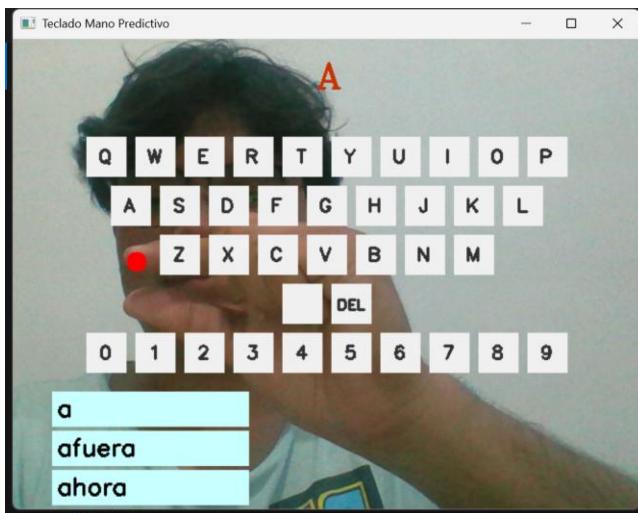
Objetivo: Evaluar si las palabras sugeridas corresponden correctamente al texto escrito y pueden ser seleccionadas sin errores.

Procedimiento:

- Se escribió el inicio de una palabra, por ejemplo: "HO".
- Se verificó si aparecían sugerencias como "HOLA", "HORARIO", "HOMBRE".
- Se seleccionó una de las sugerencias manteniendo el dedo sobre ella.

Resultado esperado: La palabra sugerida reemplaza al fragmento incompleto y se reproduce el sonido de confirmación.

Resultado obtenido: Cumplido. El sistema mostró sugerencias coherentes y facilitó la escritura fluida.



5.5 Prueba de Distribución Visual y Ergonomía

Objetivo: Confirmar que la distribución del teclado y tamaño de teclas permiten una interacción clara y sin sobreposición.

Procedimiento:

- Se revisó la posición de cada tecla en pantalla.
- Se probaron las teclas “SPACE” y “DEL” para verificar que no se empalman visualmente.
- Se evaluó la estética general del teclado.

Resultado esperado: Las teclas deben estar bien alineadas y ser fácilmente diferenciables.

Resultado obtenido: Cumplido. La versión final mejora la presentación, evita choques entre teclas especiales y mantiene una estética clara.



6. Conclusiones

El desarrollo de este prototipo demuestra cómo la inteligencia artificial puede aplicarse de forma inclusiva para facilitar la interacción humano-computadora, especialmente para personas con discapacidades motrices. A través del uso de visión por computadora con detección de manos, junto con predicción de palabras basada en trigramas, se logró construir una herramienta funcional que permite ingresar texto sin necesidad de dispositivos tradicionales como teclado o mouse.

Una de las principales fortalezas del sistema fue su capacidad para reconocer con precisión la posición del dedo índice y activar teclas mediante la permanencia, lo cual resultó ser un mecanismo intuitivo y efectivo. Además, la inclusión del gesto de palma abierta para reiniciar el texto, así como la integración de sugerencias automáticas, incrementó significativamente la usabilidad y fluidez en la escritura.

No obstante, el proyecto también presentó ciertas limitaciones:

- La precisión del sistema puede disminuir en ambientes con poca iluminación o con fondos visualmente complejos.
- La necesidad de mantener el dedo sobre una tecla por más de un segundo puede ralentizar la escritura en comparación con métodos convencionales.
- Actualmente solo se soporta un idioma (español) y un modelo de predicción relativamente básico.

Mejoras futuras propuestas:

1. **Optimización del tiempo de selección dinámica**, ajustando automáticamente el umbral según el movimiento del dedo.
2. **Adaptación a diferentes tamaños de pantalla** y resolución para mayor compatibilidad en dispositivos móviles o tablets.
3. **Entrenamiento de modelos de lenguaje más avanzados**, como redes neuronales tipo RNN o transformers para una predicción de texto más natural.
4. **Soporte multilingüe** y personalización de vocabulario para usuarios específicos.

5. **Integración con tecnologías TTS (Text to Speech)** para dar voz al texto escrito, facilitando la comunicación verbal.
6. **Inclusión de seguimiento ocular como opción adicional**, combinando métodos de control para una mayor accesibilidad.

En conclusión, este proyecto no solo cumplió con los objetivos técnicos planteados, sino que también evidenció el potencial de la inteligencia artificial como una herramienta para generar impacto social positivo. La tecnología debe estar al servicio de todos, y prototipos como este marcan el camino hacia una interacción más humana, accesible e inclusiva.