
Artículo

Realidad Virtual: SignVR

Bastian Hidalgo¹, Agustín Cárdenas², Dylan Estay³ y Matías Celedón⁴

¹ Bastián Enrique Hidalgo Covarrubias; bastian.hidalgo.c@mail.pucv.cl

² Agustín Gonzalo Cárdenas Alvarez; agustin.cardenas.a@mail.pucv.cl

³ Dylan Ignacio Estay Villalón; dylan.estay.v@mail.pucv.cl

⁴ Matías Alfredo Celedón Matus; matias.celedon.m@mail.pucv.cl

Abstract: This project proposes the development of an interactive Virtual Reality (VR) prototype aimed at teaching basic Sign Language in a didactic and engaging way. The system uses a standard webcam and computer vision techniques, particularly MediaPipe and OpenCV, to recognize hand gestures corresponding to basic signs. These recognized gestures are transmitted in real time to a Unity-based VR environment, providing immediate visual feedback and guidance to the user. The prototype targets beginners and aims to support inclusive education by offering an accessible and immersive tool for learning sign language. As a proof of concept, a "Hello World" version of the system has already been implemented, recognizing basic hand gestures such as rock, paper, and scissors. The project will evolve to include a curated set of signs and a user-friendly interface, enabling users to learn and practice through repetition, feedback, and interactive challenges. The integration of XR technologies enhances user engagement and promotes intuitive learning. This initiative combines gesture recognition, educational design, and immersive VR to create a meaningful technological solution for inclusive communication.

Keywords: Virtual Reality; Sign Language; Gesture Recognition

1. Objetivos

Desarrollar un prototipo interactivo en Realidad Virtual que permita a los usuarios aprender lenguaje de señas básico de manera didáctica, utilizando reconocimiento de gestos mediante webcam e integración con Unity y XR.

1.1. Objetivos Específicos

Para el prototipado se tienen los siguientes objetivos específicos:

- Implementar reconocimiento de señas básicas utilizando MediaPipe + OpenCV + Python.
- Integrar el reconocimiento en tiempo real con un entorno VR en Unity.
- Diseñar una interfaz intuitiva en RV que muestre feedback del gesto reconocido.
- Evaluar la factibilidad tecnológica y experiencia de usuario en el uso de esta herramienta educativa.

2. Metodología

Se usará una metodología ágil basada en Scrum, con iteraciones semanales, enfocada en la integración temprana de componentes (reconocimiento, Unity, feedback) y pruebas constantes para la validación y comprobación de las funcionalidades ya sea reconocimiento de gestos y la integración de la interfaz dinámica en respuesta al usuario.

La justificación de esta elección es por el corto periodo de tiempo en el que se estará desarrollando éste prototipo, entonces a través de esta metodología tendremos un prototipo funcional en un corto ciclo, con espacio para pruebas y retroalimentación.

3. Plan y Programa de Trabajo

A continuación se mostrará un desglose del trabajo en formato EDT que se tendrá dentro del prototipo y su desarrollo.

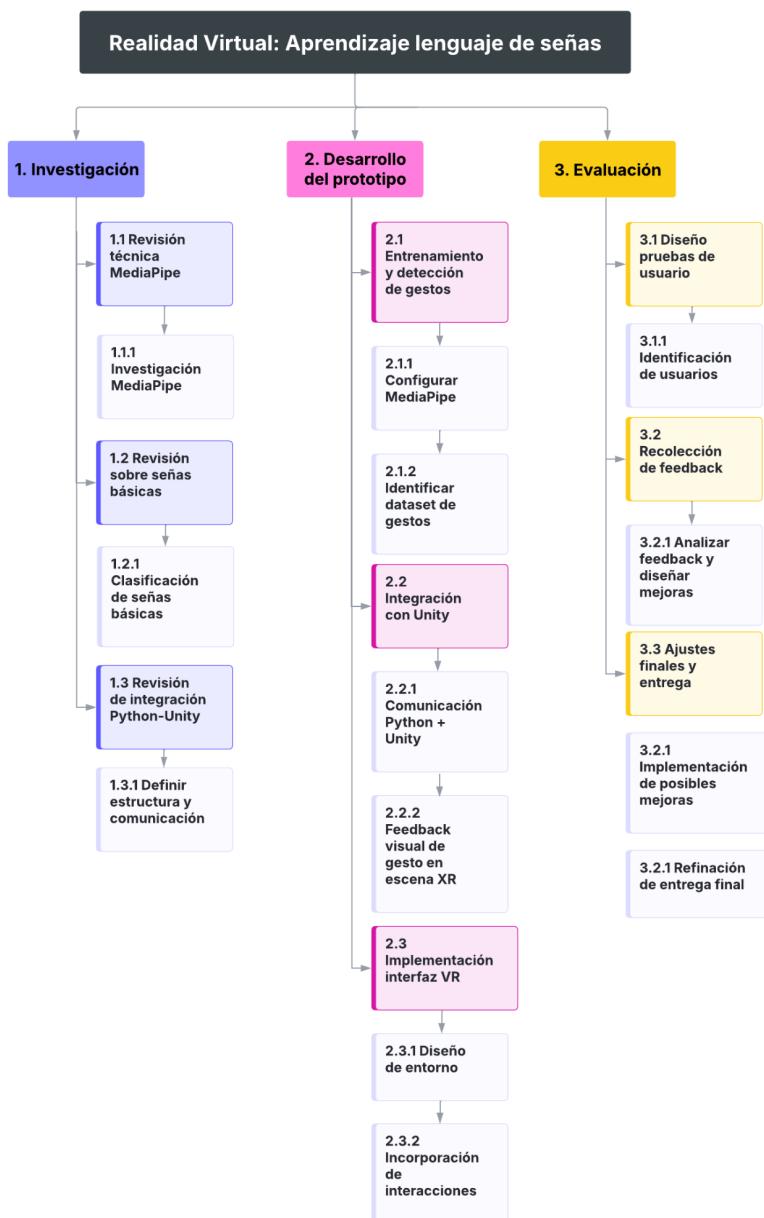


Figura 1: Desglose EDT

3.1 Programa de Trabajo

A continuación se presenta el programa de trabajo tentativo para el desarrollo del prototipo.

Semana	Fechas	Objetivo	Tareas principales	Entregables
Semana 1	4 – 10 mayo	Base técnica y teórica	- Reunión planificación - Revisión señas básicas - Estudio MediaPipe - Pruebas con webcam	Documento técnico + video corto con prueba MediaPipe
Semana 2	11 – 17 mayo	Integración Python ↔ Unity	- Clasificación básica de señas - Envío datos a Unity - Mostrar gesto reconocido	Prototipo 'Hola Mundo' funcional en Unity en base a lenguaje de señas
Semana 3	18 – 24 mayo	Ampliar dataset + interfaz VR inicial	- Agregar más señas - Mejorar precisión - Crear interfaz inicial en XR	Demo VR inicial con detección básica
Semana 4	25 – 31 mayo	Desarrollar experiencia educativa	- Diseñar entorno virtual - Agregar retroalimentación - Test interno	Video del entorno + funcionalidades educativas básicas
Semana 5	1 – 7 junio	Reconocimiento robusto + gamificación	- Refinar clasificación - Agregar puntajes o modos de prueba	Módulo educativo interactivo en XR
Semana 6	8 – 14 junio	Evaluación con usuarios	- Definir protocolo de prueba - Realizar sesiones - Recolectar feedback	Informe preliminar de resultados
Semana 7	15 – 20 junio	Ajustes finales y entrega	- Aplicar mejoras - Redactar informe - Preparar presentación	Informe final + video + entrega del prototipo

Tabla 1: Programa de Trabajo

3.2 Carta Gantt

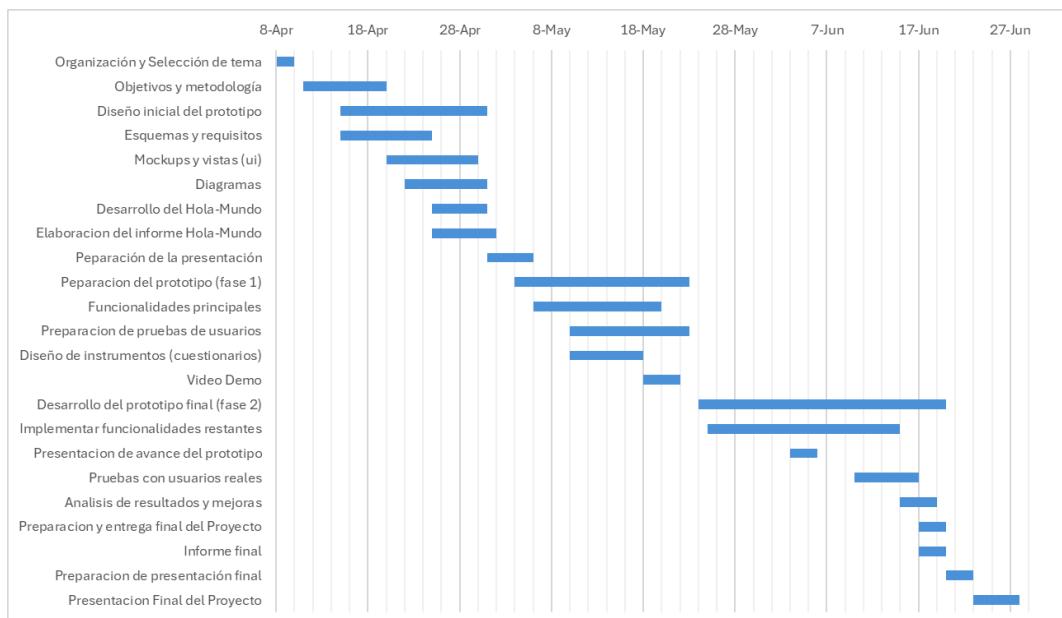


Figura 2: Carta Gantt

4. Diseño del prototipo

En esta sección se mostrará más a detalle elementos físicos y visualizables de forma tentativa en base al prototipo a desarrollar, además de una identificación de sus requisitos funcionales y no funcionales a considerar.

4.1 Esquema General

A continuación se presenta el esquema general en base a la planeación de trabajo en el desarrollo del prototipo.

1. Inicio

- El usuario se coloca el visor VR o interactúa a través de webcam, inicia el programa y accede al menú principal.
- Puede elegir: Comenzar, Calibrar, Tutorial y Opciones.

2. Selección

- El programa tendrá distintas categorías como números, letras y palabras o frases básicas.

3. Instrucción

- El sistema mostrará un modelo 3D de la señal que el usuario debe replicar.
- Instrucciones visuales guiarán al usuario.

4. Interacción

- El usuario debe intentar imitar la señal usando sus manos.
- Las cámaras del visor detectarán los gestos del usuario.
- El sistema comparará los gestos con el modelo de coincidencia.

5. Evaluación
 - Resultado correcto en caso de coincidir, se da la opción de seguir o repetir.
 - Resultado incorrecto en caso de no coincidir, el usuario debe seguir intentando hasta conseguirlo.
- .
6. Progreso
 - Se actualiza la barra de avance.
 - Al completar una lección se muestra un resumen de esta.
7. Final
 - Usuario tiene la opción de repetir, pasar a otra lección o salir.

Dibujo General de la idea:

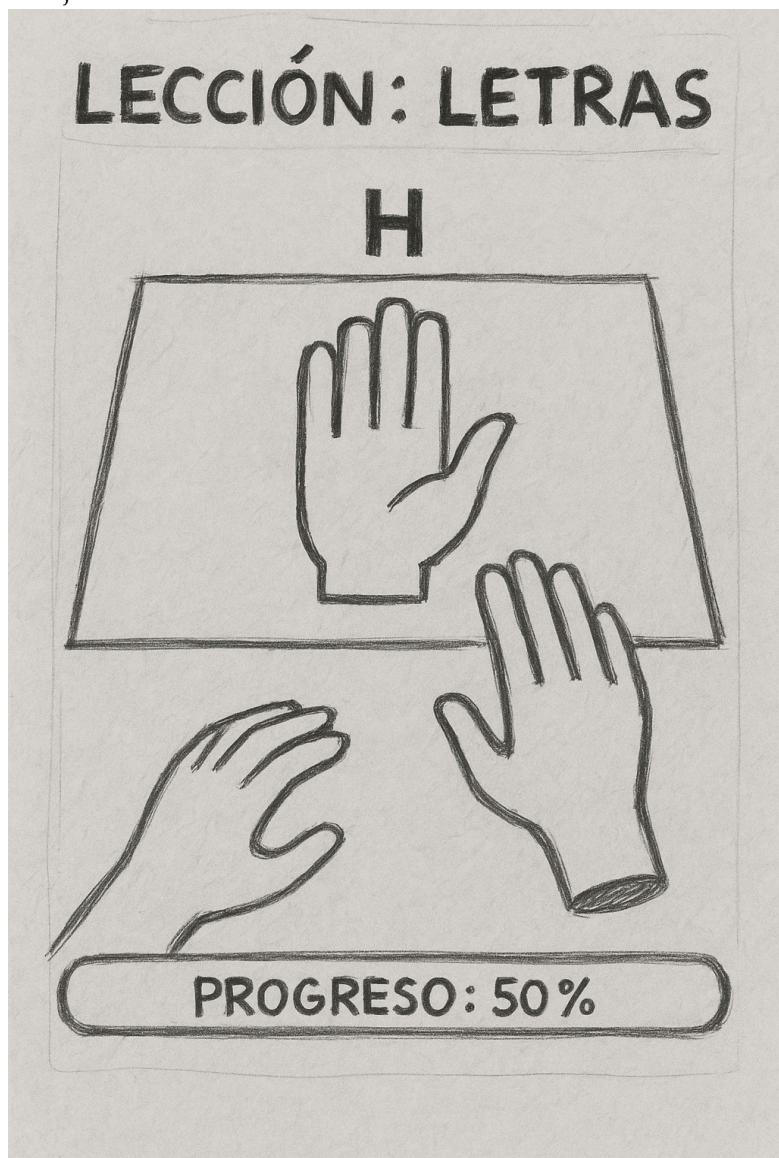


Figura 3: Sketch de la idea

4.2 Requisitos Funcionales

A continuación se presenta un listado de los requerimientos funcionales identificados dentro del prototipo del proyecto.

1. **RF01 – Detección de gestos en tiempo real:**

El sistema debe detectar y reconocer gestos básicos de lengua de señas (como el abecedario y algunas palabras comunes) usando la webcam del usuario.

2. **RF02 – Visualización del gesto reconocido:**
El gesto detectado debe ser mostrado en pantalla dentro del entorno de Unity
3. **RF03 – Integración XR (Realidad Extendida):**
El sistema debe permitir la visualización de la experiencia en un entorno XR compatible (e.g., modo VR con gafas o simulación en escritorio).
4. **RF04 – Retroalimentación educativa:**
El sistema debe ofrecer retroalimentación inmediata al usuario sobre si el gesto fue reconocido correctamente, con mensajes visuales o auditivos.
5. **RF05 – Modo aprendizaje:**
El sistema debe contar con un modo donde se enseña al usuario un gesto y este debe replicarlo con su mano frente a la cámara.
6. **RF07 – Registro de intentos:**
El sistema debe llevar un registro interno (temporal o persistente) de los gestos realizados correctamente e incorrectamente.
7. **RF07 – Comunicación entre Python y Unity:**
Debe existir una conexión estable que permita enviar datos desde Python (reconocimiento) a Unity (visualización y experiencia XR)
8. **RF08 – Calibración inicial de la cámara y manos:**
El sistema debe permitir una calibración inicial para detectar correctamente la mano y su posición.
9. **RF09 – Soporte para zurdos y diestros:**
El sistema debe permitir al usuario elegir con qué mano desea practicar los gestos.
10. **RF10 – Modo tutorial paso a paso:**
El sistema debe incluir un modo guiado paso a paso para enseñar a los nuevos usuarios cómo utilizar la aplicación.
11. **RF11 – Modo sin conexión a internet:**
El prototipo debe funcionar completamente offline, sin necesidad de conexión externa.
12. **RF12 – Modo ‘contra-reloj’:**
El sistema debe contar con un modo donde el jugador deba realizar una serie de letras en lengua de señas hasta que se acabe el tiempo.

4.3 Requisitos No Funcionales

A continuación se presenta un listado de los requerimientos no funcionales identificados dentro del prototipo del proyecto.

1. **RNF01 – Rendimiento en tiempo real:**

El sistema debe procesar los gestos y mostrar los resultados con una latencia menor a 200 ms.

2. **RNF02 – Usabilidad:**

La interfaz debe ser intuitiva y amigable, incluso para usuarios que no están familiarizados con la lengua de señas ni con XR.

3. **RNF03 – Compatibilidad de hardware:**

El sistema debe funcionar con una cámara web estándar y en computadores compatibles con Unity y XR.

4. **RNF04 – Escalabilidad de gestos:**

El sistema debe permitir la adición futura de más gestos o palabras sin necesidad de rediseñar toda la arquitectura.

5. **RNF05 – Precisión mínima aceptable:**

El sistema debe tener al menos un 85% de precisión en la detección de gestos básicos.

6. **RNF06 – Portabilidad del prototipo:**

El sistema debe poder ejecutarse en distintos dispositivos compatibles con Unity (ej: Windows o simulador VR).

7. **RNF07 – Robustez ante condiciones ambientales:**

El sistema debe funcionar de manera aceptable bajo condiciones de iluminación normales (sin necesidad de luz profesional).

8. **RNF08 – Seguridad y privacidad:**

El sistema no debe almacenar ni transmitir datos sensibles del usuario. Las imágenes de la webcam solo deben ser procesadas localmente.

9. **RNF09 – Modularidad del código:**

El sistema debe estar diseñado en módulos separados (detección, lógica, visualización) para facilitar mantenimiento y extensión.

4.4 Diagrama de Casos de Uso

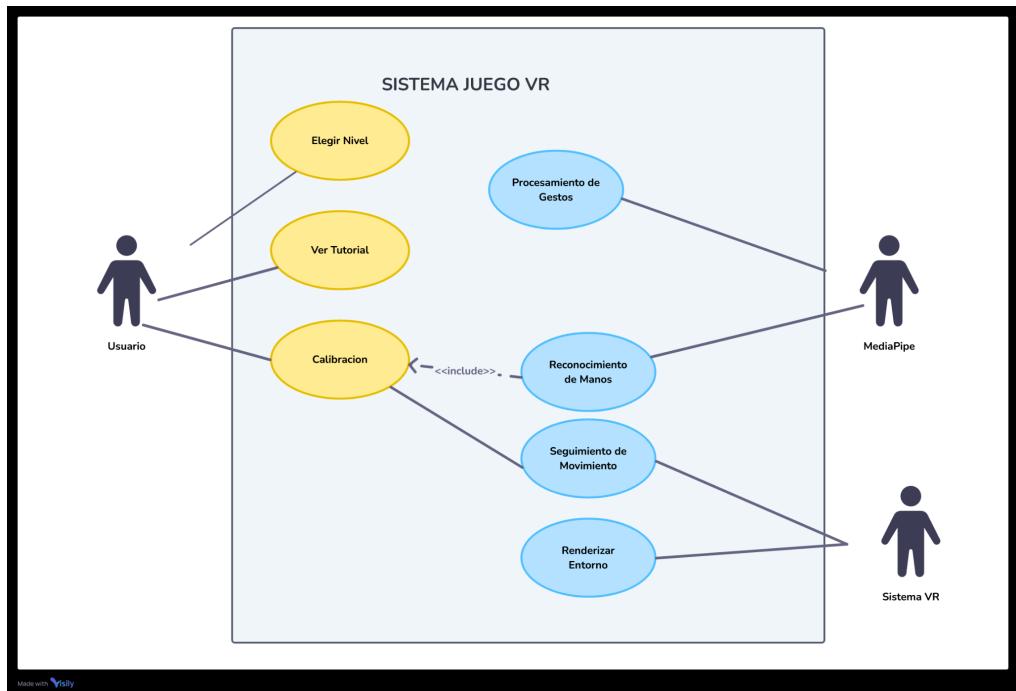


Figura 4: Caso de uso para el sistema de juego

4.5 Mockups



Figura 5: pantalla de inicio



Figura 6: pantalla

4.6 Vista Funcional

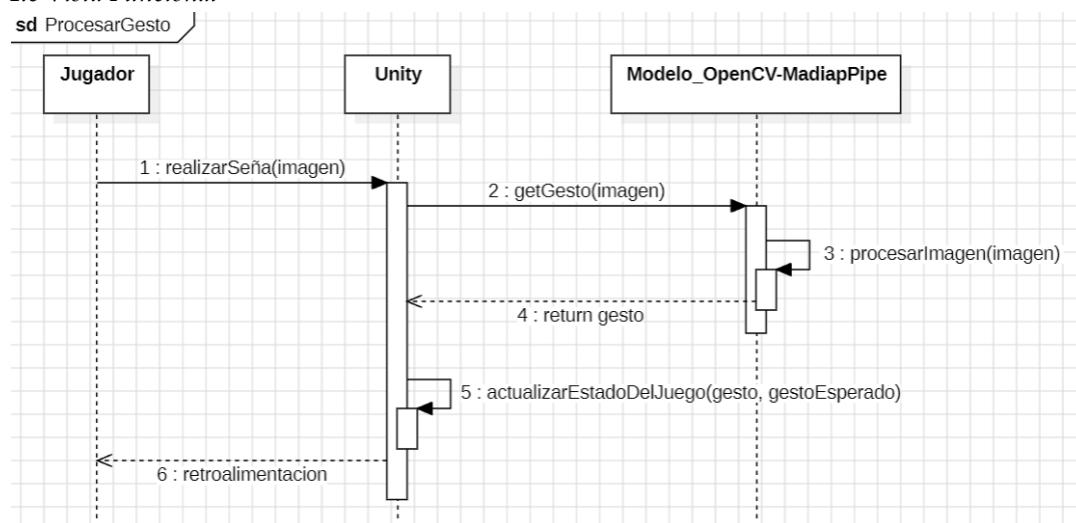


Figura 7: Diagrama de secuencia para el procesamiento de gestos

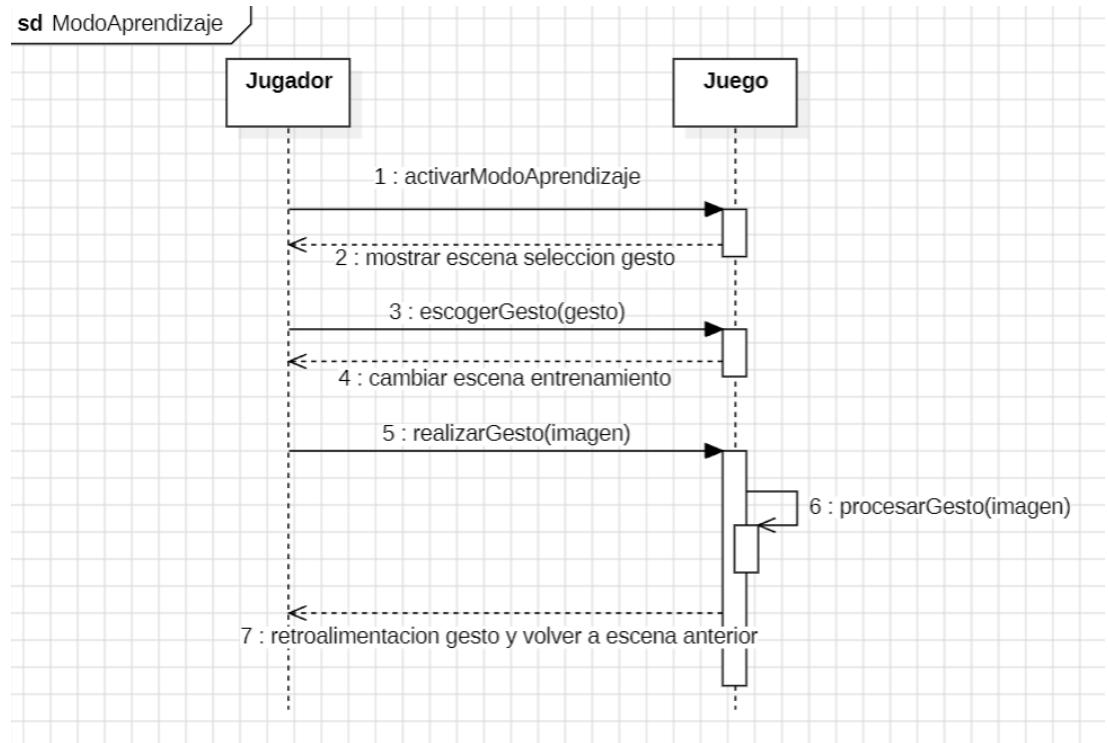


Figura 8: Diagrama de secuencia para el 'modo aprendizaje'

4.1 Vista Estática

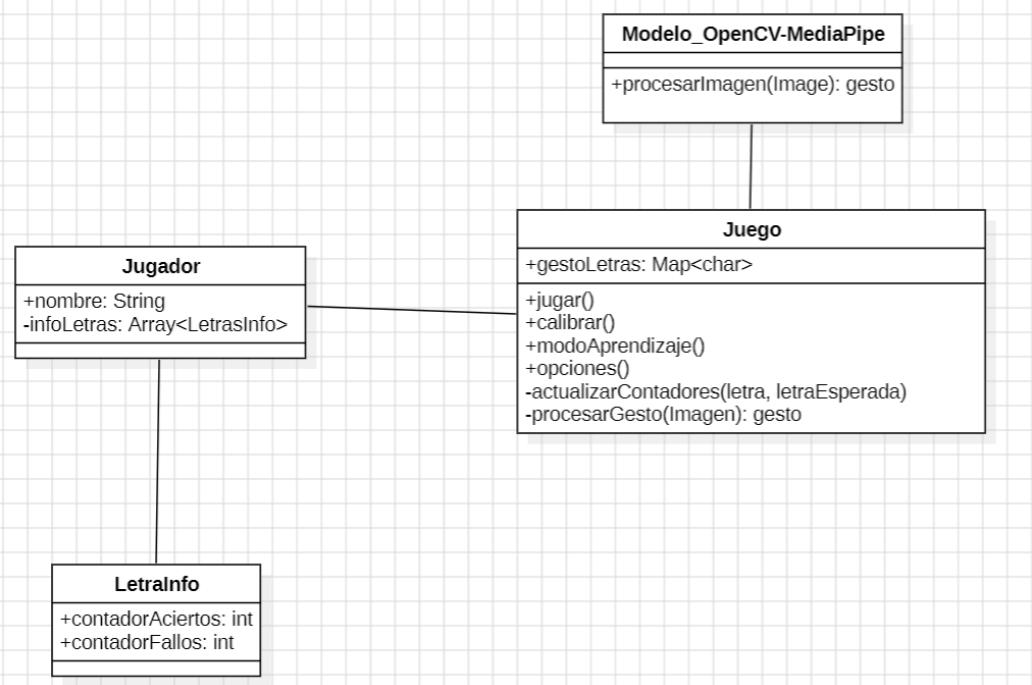


Figura 9: Diagrama de clases

4.1 Vista Dinámica

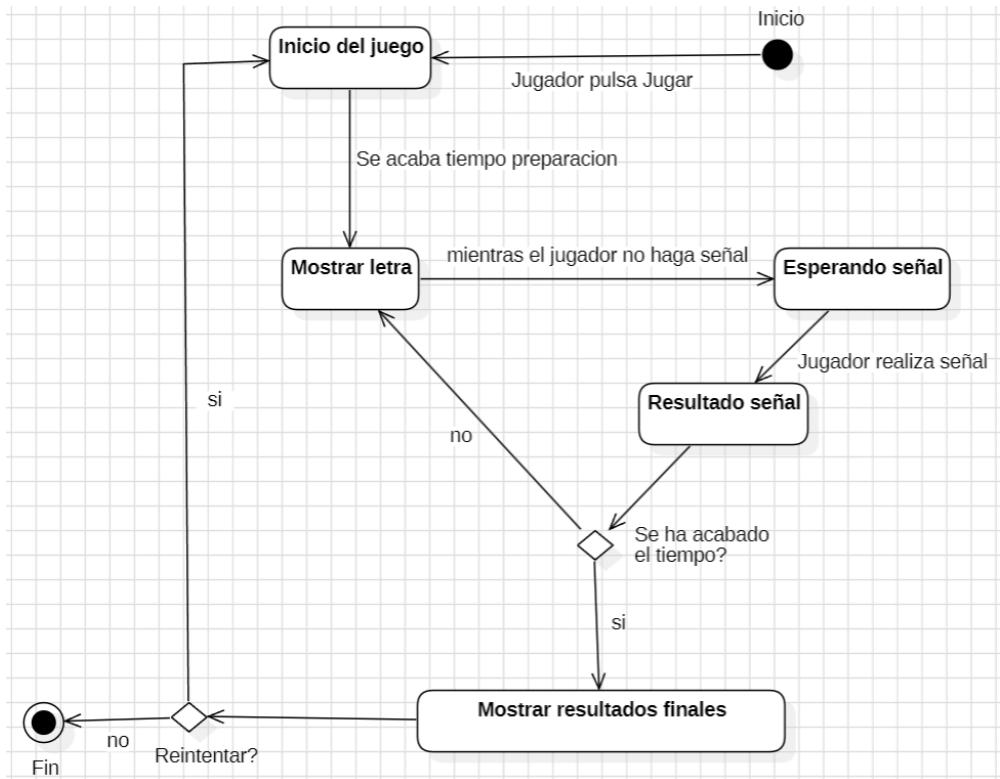


Figura 10: Diagrama de estados para el modo de juego 'contra-reloj'

5. Diseño de pruebas con usuarios

A continuación se presentará la sección donde se aborda todo el tema relacionado al diseño de la prueba con usuarios, detallando todo el proceso a realizar y tomar en cuenta.

5.1 Caracterización de los usuarios en la prueba

Para formular un correcto diseño necesitamos primero identificar qué tipo de usuarios participarán en la prueba de usuarios para definir de forma concreta los tipos de tareas y elementos a evaluar.

5.1.1 Perfil Demográfico

- **Edad:** 18-35 años (estudiantes universitarios y jóvenes profesionales)
- **Género:** Distribución equilibrada
- **Nivel Educativo:** Estudiantes universitarios o profesionales con educación superior
- **Experiencia Previa:**
 - Sin conocimiento previo en lenguaje de señas
 - Conocimiento básico en lenguaje de señas
 - Experiencia con tecnología VR

5.1.2 Criterios de selección

- **Criterios de inclusión:**
 - Estudiantes de la PUCV o profesionales jóvenes
 - Sin discapacidades visuales o motoras que impidan el uso del sistema
 - Disponibilidad para sesión de 60 minutos
 - Consentimiento informado firmado
- **Criterios de exclusión:**
 - Experiencia avanzada en lenguaje de señas
 - Problemas de salud que contraindiquen el uso de VR

5.1.3 Tamaño de muestra

- **Objetivo:** 40 participantes
- **Distribución** (tentativo):
 - 30 Estudiantes
 - 10 Jóvenes profesionales

6. Escenario y funcionalidades a probar

Hay que tener en cuenta el espacio donde se verá envuelto el usuario que probará el prototipo, además de todas las tareas a realizar en dicho ambiente y tiempo.

6.1 Escenario principal

El usuario es una persona interesada en aprender conceptos básicos de lenguaje de señas para mejorar su capacidad de comunicación inclusiva. Utilizará el prototipo VR por primera vez para comenzar su aprendizaje de forma didáctica.

6.2 Funcionalidades específicas a evaluar

- **Tarea 1:** Configuración Inicial y Tutorial
 - Colocación del visor VR
 - Navegación en menú principal
 - Completar tutorial paso a paso
 - Calibración de cámara y detección de manos
- **Tarea 2:** Modo Aprendizaje - Letras básicas
 - Acceder al modo aprendizaje
 - Aprender 5 letras del abecedario
 - Replicar gestos mostrados
 - Interpretar retroalimentación del sistema
- **Tarea 3:** Modo práctica - Palabras simples
 - Formar palabras básicas usando letras aprendidas
 - Utilizar feedback visual
 - Completar al menos 3 palabras correctamente
- **Tarea 4:** Modo Contra-reloj
 - Participar en sesión de práctica cronometrada
 - Realizar máximo número de gestos correctos en tiempo límite

7. Procedimiento de aplicación de pruebas y prueba piloto

Para una correcta formulación de las pruebas se tiene que dejar en claro la forma en la que se aplicará esta, además de su lugar, así el usuario tiene completo conocimiento de donde se debe encontrar, cómo y cuánto tiempo necesita.

7.1 Modalidad y lugar

- **Modalidad:** Presencial
- **Duración por sesión:** 60 minutos
- **Horarios:**

7.2 Protocolo de Sesión

Pre-sesión: (10 minutos)

1. Bienvenida y explicación de la prueba
2. Firma de consentimiento informado
3. Aplicación de cuestionario pre-test

Sesión principal: (30 minutos)

1. **Configuración (5 min):** Ajuste de equipo VR y calibración
2. **Tutorial (5 min):** Introducción guiada al prototipo
3. **Tareas principales (20 min):** Ejecución de las tareas programadas
4. **Observación:** Grabación de pantalla, audio y gestos del usuario

Post-sesión: (15 minutos)

1. Aplicación cuestionario post-test
2. Agradecimiento y cierre

7.3 Roles del equipo

- **Moderador principal:** Va a guiar la sesión e interactúa con el participante
- **Observadores técnicos:** Encargados de monitorear el prototipo
- **Observador UX:** Registrará los comportamientos y reacciones del participante

8. Cuestionario de medición Pre-test

8.1 Cuestionario Inicial Pre-test

Datos demográficos

1. Edad: _____ años
2. Género: Femenino Masculino No binario Prefiero no especificar

Experiencia Previa

1. **¿Ha utilizado tecnologías de realidad virtual antes?**
 Nunca Una vez Pocas veces Frecuentemente Soy experta/o
2. **¿Tiene conocimientos de lenguajes de señas?**
 Ninguno Básico Intermedio Avanzado
3. **¿Ha utilizado alguna vez experiencias educativas de este tipo?**
 Nunca Una vez Pocas veces Frecuentemente Siempre

8.1 Cuestionario Pre-test

A. System Usability Scale (SUS)

Escala Likert 1-5 (1=Totalmente en desacuerdo, 5=Totalmente de acuerdo)

1. Creo que me gustará usar este sistema VR frecuentemente
2. Encontraré el sistema innecesariamente complejo
3. Pienso que el sistema será fácil de usar
4. Creo que necesitaré el apoyo de una persona técnica para poder usar este sistema
5. Encontraré que las diversas funciones en este sistema estarán bien integradas
6. Me parece que habrá demasiada inconsistencia en este sistema.
7. Imagino que la mayoría de las personas aprenderán a usar este sistema muy rápidamente.
8. Encontraré el sistema muy engorroso de usar.
9. Me sentiré muy segura/o al usar el sistema.
10. Necesitaré aprender muchas cosas antes de poder usar el sistema.

B. Intrinsic Motivation Inventory (IMI)

Escala Likert 1-5 (1=Totalmente en desacuerdo, 5=Totalmente de acuerdo)

Interés/Disfrute:

1. Disfrutaré mucho usando este sistema VR
2. Esta actividad será divertida para mí
3. Encontraré esta actividad muy interesante
4. Esta actividad captará mi atención en absoluto
5. Describiré esta actividad como muy placentera

Competencia Percibida:

1. Creo que seré bastante buena/o en esta actividad
2. Estaré satisfecha/o con mi desempeño en esta tarea
3. Después de trabajar en esta actividad por un tiempo, me sentiré bastante competente
4. Estaré muy satisfecha/o con mi desempeño en esta actividad
5. Me sentiré como si fuera bastante efectiva/o completando esta actividad

Presión/Tensión:

1. No me sentiré nerviosa/o para nada mientras haga esto
2. No me sentiré muy tensa/o mientras haga esta actividad
3. No me sentiré presionada/o mientras haga esto
4. Me sentiré muy relajada/o haciendo esto
5. No estaré ansiosa/o mientras trabajo en esta tarea

C. Holistic Multi-modal Skill Acquisition Model (HMSAM)

Escala Likert 1-5 (1=Totalmente en desacuerdo, 5=Totalmente de acuerdo)

Inmersión:

1. Me sentiré completamente inmersa/o en el entorno virtual
2. Perderé la noción del tiempo mientras uso el sistema
3. Me olvidaré del mundo real mientras estoy en la experiencia VR
4. Me sentiré como si realmente estuviera en el entorno virtual
5. La experiencia me mantendrá completamente absorta/o

8.2 Cuestionario Post-test

A. System Usability Scale (SUS)

Escala Likert 1-5 (1=Totalmente en desacuerdo, 5=Totalmente de acuerdo)

1. Creo que me gustaría usar este sistema VR frecuentemente
2. Encontré el sistema innecesariamente complejo
3. Pensé que el sistema era fácil de usar
4. Creo que necesitaría el apoyo de una persona técnica para poder usar este sistema
5. Encontré que las diversas funciones en este sistema estaban bien integradas
6. Me parece que había demasiada inconsistencia en este sistema.
7. Imagino que la mayoría de las personas aprenderían a usar este sistema muy rápidamente.
8. Encontré el sistema muy engorroso de usar.
9. Me sentí muy seguro al usar el sistema.
10. Necesité aprender muchas cosas antes de poder usar el sistema.

B. Intrinsic Motivation Inventory (IMI)

Escala Likert 1-5 (1=Totalmente en desacuerdo, 5=Totalmente de acuerdo)

Interés/Disfrute:

1. Disfruté mucho usando este sistema VR
2. Esta actividad fue divertida para mí
3. Encontré esta actividad muy interesante
4. Esta actividad captó mi atención en absoluto
5. Describiría esta actividad como muy placentera

Competencia Percibida:

1. Creo que soy bastante bueno/a en esta actividad
2. Estoy satisfecho/a con mi desempeño en esta tarea
3. Después de trabajar en esta actividad por un tiempo, me sentí bastante competente
4. Estoy muy satisfecho/a con mi desempeño en esta actividad
5. Me sentí como si fuera bastante efectivo/a completando esta actividad

Presión/Tensión:

1. No me sentí nervioso/a para nada mientras hacía esto
2. No me sentí muy tenso/a mientras hacía esta actividad
3. No me sentí presionado/a mientras hacía esto
4. Me sentí muy relajado/a haciendo esto
5. No estuve ansioso/a mientras trabajaba en esta tarea

C. Holistic Multi-modal Skill Acquisition Model (HMSAM)

Escala Likert 1-5 (1=Totalmente en desacuerdo, 5=Totalmente de acuerdo)

Inmersión:

1. Me sentí completamente inmerso/a en el entorno virtual
2. Perdí la noción del tiempo mientras usaba el sistema
3. Me olvidé del mundo real mientras estaba en la experiencia VR
4. Me sentí como si realmente estuviera en el entorno virtual
5. La experiencia me mantuvo completamente absorto/a

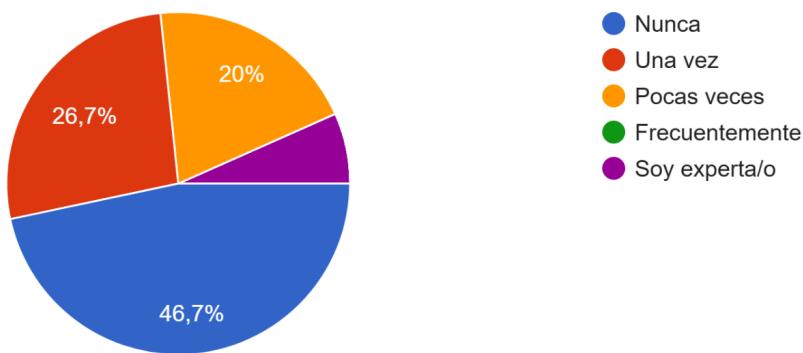
8.3 Cuestionario de opinión

1. ¿Encontró el sistema innecesariamente complejo? ¿Por qué?
2. ¿Describiría esta actividad como placentera? ¿Por qué?
3. Después de trabajar en esta actividad, ¿Se sintió competente? ¿Por qué?
4. ¿Se sintió relajada/o durante la actividad? ¿Por qué?
5. ¿Se sintió durante la actividad si realmente estuviera en el entorno virtual? ¿Por qué?

9. Análisis de resultados

¿Ha utilizado tecnologías de realidad virtual antes?

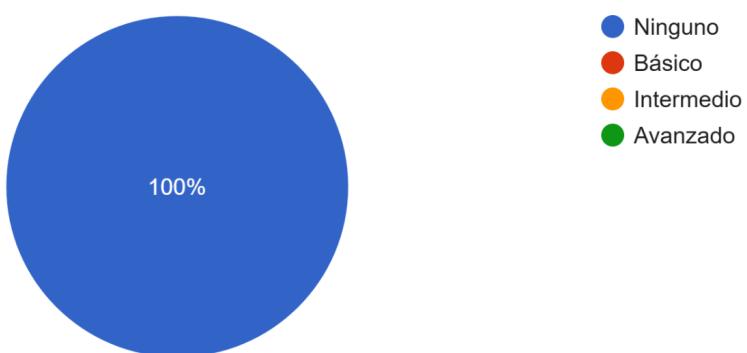
15 respuestas



Se logró reunir una cantidad equitativa entre personas que saben utilizar tecnologías de realidad virtual con personas que no tenían ningún conocimiento.

¿Tiene conocimientos de lenguaje de señas?

15 respuestas



Ninguno poseía conocimiento en Lenguaje de señas, permitiendo así que hubiera una mayor oportunidad de aprendizaje y motivación.

En general en base a las respuestas se tuvo una expectativa mediana entre 2 a 3 en la escala Likert, y luego en la realidad hubo un leve aumento en la escala 3 a 4 Likert, esto puede ser debido al desconocimiento que tenían a simple vista y luego a través de la experiencia y la actividad se logró mejorar la visión que tenían de la actividad y su desempeño propio.