



UNIVERSIDAD TÉCNICA DE AMBATO

**FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E
INDUSTRIAL**

**CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES E
INFORMÁTICOS**

Tema:

**PLATAFORMA DE INTERACCIÓN VIRTUAL PARA LA
NAVEGACIÓN DE VISITANTES PRIMARIOS EN LOS PREDIOS
DE LA UTA DEL CAMPUS HUACHI USANDO UN AVATAR Y
ALGORITMOS DE MACHINE LEARNING.**

Trabajo de Titulación Modalidad: Proyecto de Investigación, presentado
previo a la obtención del título de Ingeniero en Sistemas Computacionales e
Informáticos

ÁREA: Inteligencia artificial

LÍNEA DE INVESTIGACIÓN: Tecnologías de la Información

AUTOR: Alexis Israel Palate Amaguaña

TUTOR: Ing. Rubén Nogales, Mg.

Ambato – Ecuador

agosto – 2023

APROBACIÓN DEL TUTOR

En calidad de tutor del Trabajo de Titulación con el tema: PLATAFORMA DE INTERACCIÓN VIRTUAL PARA LA NAVEGACIÓN DE VISITANTES PRIMARIOS EN LOS PREDIOS DE LA UTA DEL CAMPUS HUACHI USANDO UN AVATAR Y ALGORITMOS DE MACHINE LEARNING, desarrollado bajo la modalidad Proyecto de Investigación por el señor Alexis Israel Palate Amaguaña, estudiante de la Carrera de Ingeniería en Sistemas Computacionales e Informáticos, de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial, de la Universidad Técnica de Ambato, me permito indicar que el estudiante ha sido tutorado durante todo el desarrollo del trabajo hasta su conclusión, de acuerdo a lo dispuesto en el Artículo 15 del Reglamento para obtener el Título de Tercer Nivel, de Grado de la Universidad Técnica de Ambato, y el numeral 7.4 del respectivo instructivo.

Ambato, agosto 2023.

Ing. Rubén Nogales, Mg.

TUTOR

AUTORÍA

El presente Proyecto de Investigación titulado: PLATAFORMA DE INTERACCIÓN VIRTUAL PARA LA NAVEGACIÓN DE VISITANTES PRIMARIOS EN LOS PREDIOS DE LA UTA DEL CAMPUS HUACHI USANDO UN AVATAR Y ALGORITMOS DE MACHINE LEARNING es absolutamente original, auténtico y personal. En tal virtud, el contenido, efectos legales y académicos que se desprenden del mismo son de exclusiva responsabilidad del autor.

Ambato, agosto 2023.



Alexis Israel Palate Amaguaña

C.C. 1804390027

AUTOR

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que haga uso de este Trabajo de Titulación como un documento disponible para la lectura, consulta y procesos de investigación.

Cedo los derechos de mi Trabajo de Titulación en favor de la Universidad Técnica de Ambato, con fines de difusión pública. Además, autorizo su reproducción total o parcial dentro de las regulaciones de la institución.

Ambato, agosto 2023.



Alexis Israel Palate Amaguaña

C.C. 1804390027

AUTOR

APROBACIÓN DEL TRIBUNAL DE GRADO

En calidad de par calificador del Informe Final del Trabajo de Titulación presentado por el señor Alexis Israel Palate Amaguaña, estudiante de la Carrera de Ingeniería en Sistemas Computacionales e Informáticos, de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial, bajo la Modalidad Proyecto de Investigación, titulado PLATAFORMA DE INTERACCIÓN VIRTUAL PARA LA NAVEGACIÓN DE VISITANTES PRIMARIOS EN LOS PREDIOS DE LA UTA DEL CAMPUS HUACHI USANDO UN AVATAR Y ALGORITMOS DE MACHINE LEARNING, nos permitimos informar que el trabajo ha sido revisado y calificado de acuerdo al Artículo 17 del Reglamento para obtener el Título de Tercer Nivel, de Grado de la Universidad Técnica de Ambato, y al numeral 7.6 del respectivo instructivo. Para cuya constancia suscribimos, conjuntamente con la señora Presidente del Tribunal.

Ambato, agosto 2023.

Ing. Pilar Urrutia, Mg.

PRESIDENTE DEL TRIBUNAL

Ing. Hernando Buenaño, Mg.
PROFESOR CALIFICADOR

Ing. Carlos Nuñez, Mg.
PROFESOR CALIFICADOR

DEDICATORIA

A Dios, por la vida, la salud, y la sabiduría para tomar las decisiones correctas a lo largo de mi vida.

A mi familia, por la disposición de apoyarme en todo momento.

A mis amigos, por el apoyo incondicional brindado en los momentos requeridos.

A mi tutor, por su guía, su apoyo, sus conocimientos y sus consejos durante todo este proceso de formación profesional.

AGRADECIMIENTO

Agradezco en primer lugar a Dios, por darme la fortaleza de no desistir a pesar de las adversidades que se presentaron a lo largo de mi vida estudiantil.

A mis mejores amigos, Estefanía, Steven y Diego, por ser la perfecta definición de lo que yo considero “amistad”. Por el apoyo en momentos cruciales de mi vida, por sus consejos, y, sobre todo, por su tiempo.

A mi familia, que, a pesar de todo, siempre se mantuvo unida demostrando la inquebrantable relación que formó Dios a lo largo de nuestra vida.

A todos los compañeros que se vieron envueltos en las decisiones que tomé durante mi proceso académico con el fin de cumplir esta meta. Gracias por formar parte de esta grandiosa etapa de mi vida.

A aquella persona que no estuvo al inicio de mi carrera, pero sí al final. Una clara, breve y perfecta demostración de que no es el tiempo, si no, la persona. Gracias por su apoyo en el momento más complicado de mi vida.

Un agradecimiento profundo y sincero a todos los docentes que me brindaron sus conocimientos durante mi vida universitaria. Especialmente al Ing. Rubén Nogales, que además de su apoyo y sus conocimientos como tutor, me supo demostrar sus valores y destrezas como profesional. Sin su ayuda, nada de esto hubiera sido posible.

ÍNDICE GENERAL DE CONTENIDOS

RESUMEN EJECUTIVO	xiii
ABSTRACT	xiv
CAPITULO I: MARCO TEÓRICO.....	1
1.1. Tema de investigación.....	1
1.2. Antecedentes investigativos	1
1.2.1. Contextualización del problema.....	1
1.2.2. Antecedentes Investigativos.....	3
1.2.3. Fundamentación Teórica.....	11
1.3. Objetivos	14
1.3.1. Objetivo General	14
1.3.2. Objetivos Específicos.....	14
CAPÍTULO II	15
METODOLOGÍA	15
2.1. Materiales.....	15
2.1.1. Humanos	15
2.1.2. Institucionales	15
2.1.3. Otros.....	15
2.2. Métodos.....	16
2.2.1. Modalidad de la Investigación	16
2.2.2. Población y muestra	16

2.2.3. Recolección de la información.....	17
CAPÍTULO III.....	19
RESULTADOS Y DISCUSIÓN.....	19
3.1. Desarrollo de la propuesta.....	19
3.1.1. Diseño de una plataforma de interacción virtual que permita reconocer gestos estáticos de la mano	19
3.1.2. Diseño de una plataforma de interacción virtual que permita reconocer gestos estáticos de la mano para la manipulación del avatar y su interacción con el ambiente	21
3.1.3. Implementación de la plataforma de interacción virtual para la navegación en los predios de la UTA del campus Huachi usando un avatar y algoritmos de Machine Learning	24
3.2. Experimentación	26
3.2.1. Primera parte	26
3.2.2. Segunda parte.....	30
3.3. Resultados de las pruebas con la plataforma.....	33
CAPÍTULO IV.....	45
CONCLUSIONES Y RECOMENDACIONES.....	45
4.1. Conclusiones	45
4.2. Recomendaciones.....	46
BIBLIOGRAFÍA.....	47
Anexos.....	50
Anexo 1. Primera parte de la encuesta	50

Anexo 2. Segunda parte de la encuesta.....	51
--	----

ÍNDICE DE TABLAS

Tabla 1. Revisión Sistemática de la Literatura.....	5
Tabla 2. Criterios de Inclusión y Exclusión	6
Tabla 3. Aspectos y Técnicas Comprendidas por la IA.....	12
Tabla 4. Recursos Económicos	15
Tabla 5. Técnica de Recolección de Información.....	18
Tabla 6. Acciones realizadas por el avatar.....	22
Tabla 7. Características del computador	27
Tabla 8. Resultados de la primera prueba del modelo	27
Tabla 9. Resultados de la segunda prueba del modelo.....	28
Tabla 10. Resultados de la tercera prueba del modelo.....	29
Tabla 11. Resultados de la cuarta prueba del modelo	29
Tabla 12. Resultados de la quinta prueba del modelo.....	30
Tabla 13. Resultados de la primera parte de la encuesta.....	35
Tabla 14. Resultados de la segunda parte de la encuesta.....	41
Tabla 15. Respuestas - Pregunta 10	44

ÍNDICE DE FIGURAS

Figura 1. Esquema del modelo de reconocimiento de gestos de la mano	20
Figura 2. Modelo del avatar	23
Figura 3. Capsule Collider aplicado al Avatar	24
Figura 4. Integración de los IDE	25
Figura 5. Clasificación de KNN con dos clases (0 y 1)	26
Figura 6. Modelo 3D generado en Infracworks	31
Figura 7. Modelo 3D generado en Sketchup.....	32
Figura 8. Esquema de las pruebas realizadas	34
Figura 9. Respuestas - Pregunta 2.....	35
Figura 10. Respuestas - Pregunta 3.....	36
Figura 11. Respuestas - Pregunta 4.....	37
Figura 12. Respuestas - Pregunta 5.....	38
Figura 13. Respuestas - Pregunta 6.....	39
Figura 14. Respuestas - Pregunta 7.....	41
Figura 15. Respuestas - Pregunta 8.....	42
Figura 16. Respuestas - Pregunta 9.....	43

RESUMEN EJECUTIVO

Las personas que no conocen un lugar de su interés tienen problemas en su primera visita, produciéndoles una mala experiencia. En este contexto, una persona puede buscar alternativas para conocer estos lugares de forma virtual, antes de conocerlos físicamente. Sin embargo, la mayoría de las opciones no son un reflejo de la realidad. Limitándolos básicamente a ver los exteriores de un lugar. En este sentido, se propone una plataforma de interacción virtual donde el usuario puede controlar un avatar con los gestos de la mano, para conocer un lugar de forma virtual antes de conocerlo físicamente. De esta forma puede mejorar su experiencia en la primera visita a los lugares que son de su interés. En este proyecto se propone un modelo 3D del campus Huachi de la Universidad Técnica de Ambato (UTA). Esta plataforma funciona con dos sistemas integrados, ejecutándose sobre el mismo computador. El primer sistema consiste en un modelo de reconocimiento de gestos de la mano. Este sistema utiliza el sensor Leap Motion Controller para capturar los gestos realizados por el usuario. Estos datos son enviados hacia el modelo. Este se encarga de procesarlos utilizando un clasificador KNN, para predecir una etiqueta que representa el gesto. Posteriormente, se envía la etiqueta predicha a través del protocolo de comunicación TCP/IP, con la dirección IP local del computador donde se ejecuta la plataforma. El segundo sistema recibe la etiqueta y lo interpreta como una instrucción que el avatar debe cumplir realizando una acción dentro de su entorno virtual. La plataforma fue testeada por 15 personas. Un grupo de 5 personas que sí conocen la UTA, y otro grupo de 10 personas que no conocen la UTA, dando prioridad a las personas que no conocen la UTA. El modelo fue testeado por el investigador, realizando 20 repeticiones por cada gesto. Encontrando que, con $kn=5$ y $frames=400$, se obtiene un porcentaje de exactitud de 100% en el gesto Open Hand, 85% en el gesto Close Hand, 65% en el gesto Wave In, 95% en el gesto Wave Out, y 10% en el gesto Pinch. Estos gestos se utilizan para la interacción del avatar con el mundo virtual. Sin embargo, se limita la acción correspondiente al gesto Pinch, debido a su bajo porcentaje de exactitud.

Palabras Clave: Plataforma de interacción virtual, Avatar, reconocimiento de gestos de la mano, Leap Motion Controller, KNN, TCP/IP, Unity 3D, Sketchup, Infracore.

ABSTRACT

People who do not know a place of interest have problems on their first visit, causing them a bad experience. In this context, a person can look for alternatives to get to know these places virtually, before getting to know them physically. However, most of the options are not a reflection of reality. Basically limiting them to see the exteriors of a place. In this sense, a virtual interaction platform is proposed where the user can control an avatar with hand gestures, to get to know a place virtually before getting to know it physically. In this way you can improve your experience on the first visit to the places that are of interest to you. In this project, a 3D model of the Huachi campus of the Technical University of Ambato (UTA) is proposed. This platform works with two integrated systems, running on the same computer. The first system consists of a hand gesture recognition model. This system uses the Leap Motion Controller sensor to capture the gestures made by the user. These data are sent to the model. This is in charge of processing them using a KNN classifier, to predict a label that represents the gesture. Subsequently, the predicted label is sent through the TCP/IP communication protocol, with the local IP address of the computer where the platform is running. The second system receives the label and interprets it as an instruction that the avatar must comply with by performing an action within its virtual environment. The platform was tested by 15 people. A group of 5 people who do know the UTA, and another group of 10 people who do not know the UTA, giving priority to people who do not know the UTA. The model was tested by the researcher, performing 20 repetitions for each gesture. Finding that, with $kn=5$ and $frames=400$, an accuracy percentage of 100% is obtained in the Open Hand gesture, 85% in the Close Hand gesture, 65% in the Wave In gesture, 95% in the Wave Out gesture, and 10% on the Pinch gesture. These gestures are used for the interaction of the avatar with the virtual world. However, the action corresponding to the Pinch gesture is limited, due to its low percentage of accuracy.

Keywords: Virtual interaction platform, Avatar, hand gesture recognition, Leap Motion Controller, KNN, TCP/IP, Unity 3D, Sketchup, Infracworks.

CAPITULO I: MARCO TEÓRICO

1.1. Tema de investigación

PLATAFORMA DE INTERACCIÓN VIRTUAL PARA LA NAVEGACIÓN DE VISITANTES PRIMARIOS EN LOS PREDIOS DE LA UTA DEL CAMPUS HUACHI USANDO UN AVATAR Y ALGORITMOS DE MACHINE LEARNING.

1.2. Antecedentes investigativos

1.2.1. Contextualización del problema

Los Tours Virtuales (VT) son una de las mejores opciones cuando la visita a un lugar no se puede realizar en persona o es mejor evitarla por riesgos como enfermedades, desastres naturales o inseguridad [1]. Un claro ejemplo es lo ocurrido durante el mes de marzo de 2020 donde la Organización Mundial de la Salud (OMS) determinó que la COVID-19 se puede caracterizar como una pandemia a nivel mundial [2]. Debido a esto, y para evitar contagios, gran parte de la población del Ecuador tuvo que aislarse totalmente perdiendo el acceso físico a sus lugares de trabajo.

Del mismo modo, en el área de la educación se evidenciaron problemas con los estudiantes universitarios de nuevo ingreso. Estos accedieron a la universidad durante la pandemia ocasionando que en sus primeros años de vida estudiantil no conocieran el establecimiento educativo de forma física, pues desempeñaban sus labores estudiantiles mediante clases virtuales. El problema consiste en que los estudiantes no conocen los sitios relevantes dentro de la institución. Esto provoca que tengan una mala experiencia en su primera visita a la institución. Por tal motivo, se propone una alternativa para conocer estos lugares de forma virtual antes de conocerlos físicamente. Esta alternativa consiste en una plataforma de navegación virtual basada en el reconocimiento de gestos de la mano.

En este contexto, los VT pueden definirse como una representación de la realidad, donde el uso de la Realidad Virtual (VR), es decir, imágenes captadas de la realidad o generadas de forma sintética brindan al usuario la posibilidad de navegar e interactuar con su entorno [1]. La VR puede ser experimentada por el usuario mediante el uso de cascos o auriculares 3D como Facebook Oculus, HTC Vive, Samsung Gear VR, Google Card Board o mediante la pantalla de un computador [3]. Además, es necesario el uso de sensores que capten los gestos realizados por el usuario. Estos dispositivos pueden clasificarse en sensores sin contacto y sensores de contacto con el usuario. Los sensores sin contacto utilizan una o varias cámaras. Mientras que los sensores de contacto usan guantes instrumentados o brazaletes. Entre los dispositivos de contacto se mencionan a CyberGlove, Myo Armband, Wiimote. Y entre los dispositivos sin contacto se encuentran: Cámaras de Video, Cámaras estéreo, Kinect y Leap Motion Controller (LMC) [4][5].

En este sentido, la Plataforma de Navegación Virtual (PNV) propuesta pretende ayudar a los estudiantes que visitan la Universidad Técnica de Ambato (UTA) por primera vez reduciendo la mala experiencia que tendrían en su primera visita a la institución. Esta PNV está basada en el reconocimiento de gestos de la mano mediante el sensor LMC pues captura la posición de los dedos con una precisión muy alta [6]. Estos gestos pueden ser estáticos o dinámicos. Donde un gesto estático consiste en una posición de la mano bien definida en un momento dado. Mientras que un gesto dinámico es una secuencia de movimientos durante un lapso de tiempo [7]. Entre los gestos estáticos que reconoce el LMC se encuentran: OPEN HAND, CLOSE HAND, WAVE IN, WAVE OUT y PINCH. Estos gestos son interpretados por un modelo de Machine Learning y retorna una acción en respuesta al gesto realizado por el usuario. Del mismo modo, permite a los usuarios navegar e interactuar en un entorno virtual similar a la institución antes de conocerlo físicamente. De esta forma, el usuario tendrá la posibilidad de controlar un avatar y navegar virtualmente a través de los predios de la UTA realizando acciones como: abrir y cerrar puertas y ventanas, girar hacia la izquierda y derecha, encender la luz o simplemente caminar por el entorno virtual.

Línea de investigación: Tecnologías de la información.

Sub líneas de investigación: Procesamiento de la información.

1.2.2. Antecedentes Investigativos

Con el propósito de obtener información relevante acerca del tema de este proyecto, se definieron palabras clave como Leap Motion Controller, Virtual Worlds, Unity 3D, KNN, Avatar, y Virtual Tours para realizar una investigación exhaustiva dentro del período 2016 – 2022, encontrando y destacando los siguientes trabajos:

	Revista	Cadena	Número de artículos
Cadena 1	IEEE Xplore	"All Metadata":Leap Motion Controller with KNN	1
	SCOPUS	TITLE-ABS-KEY (leap AND motion AND controller AND with AND kcn) AND PUBYEAR > 2015 AND PUBYEAR < 2023	5
	SPRINGER Link	'Leap AND Motion AND Controller AND with AND KNN' within 2016 - 2022	726
	ACM Digital Library	[All: leap motion controller with knn] AND [All: leap motion controller] AND [All: knn] AND [E-Publication Date: (01/01/2016 TO 12/31/2022)]	3.691
Cadena 2	IEEE Xplore	("All Metadata":Leap Motion Controller in Unity 3D)	8
	SCOPUS	TITLE-ABS-KEY (leap AND motion AND controller AND in AND unity AND 3d) AND PUBYEAR > 2015 AND PUBYEAR < 2023	18
	SPRINGER Link	'Leap AND Motion AND Controller AND in AND Unity AND 3D'	1.464

		within 2016 - 2022	
	ACM Digital Library	[All: leap motion controller in unity 3d] AND [All: leap motion controller] AND [All: knn] AND [All: unity] AND [E-Publication Date: (01/01/2016 TO 12/31/2022)]	721
Cadena 3	IEEE Xplore	“Leap Motion Controller in Virtual Worlds”	9
	SCOPUS	TITLE-ABS-KEY (leap AND motion AND controller AND in AND virtual AND worlds) AND PUBYEAR > 2015 AND PUBYEAR < 2023	37
	SPRINGER Link	'Leap AND Motion AND Controller AND in AND Virtual AND Worlds AND "Leap Motion Controller" AND (Virtual OR World)' within 2016 - 2022	387
	ACM Digital Library	[All: leap motion controller in virtual worlds] AND [All: leap motion controller] AND [All: knn] AND [All: virtual world] AND [E-Publication Date: (01/01/2016 TO 12/31/2022)]	2.427
Cadena 4	IEEE Xplore	("All Metadata": Controlling an avatar with Leap Motion Controller)	2
	SCOPUS	TITLE-ABS-KEY (controlling AND an AND avatar AND with AND leap AND motion AND controller) AND PUBYEAR > 2015 AND PUBYEAR > 2015	1
	SPRINGER Link	'Controlling AND an AND avatar AND with AND Leap AND Motion AND Controller AND "Leap Motion	144

		Controller" AND (avatar)' within 2016 - 2022	
	ACM Digital Library	[All: controlling an avatar with leap motion controller] AND [All: leap motion controller] AND [All: knn] AND [All: avatar] AND [E-Publication Date: (01/01/2016 TO 12/31/2022)]	72
Cadena 5	IEEE Xplore	("All Metadata":control of an avatar in unity 3d) OR ("All Metadata":Leap Motion Controller) AND ("All Metadata":Avatar) AND ("All Metadata":Unity 3D) Filters Applied: 2016 - 2022	8
	SCOPUS	(TITLE-ABS-KEY (control AND of AND an AND avatar AND in AND unity AND 3d) OR TITLE-ABS-KEY (leap AND motion AND controller) AND TITLE-ABS-KEY (avatar) OR TITLE-ABS-KEY (unity AND 3d)) AND PUBYEAR > 2015 AND PUBYEAR < 2023	32
	SPRINGER Link	'control AND of AND an AND avatar AND in AND unity AND 3d AND "Avatar" AND (unity OR 3D)' within 2016 - 2022	415
	ACM Digital Library	[All: control of an avatar in unity 3d] AND [All: leap motion controller] AND [All: avatar] AND [All: unity 3d] AND [E-Publication Date: (01/01/2016 TO 12/31/2022)]	2.716
Total:			12.884

Tabla 1. Revisión Sistemática de la Literatura

Elaborado por: El Investigador

De acuerdo con la revisión sistemática de la literatura, existe un total de 12.884 artículos relacionados al trabajo propuesto en este proyecto. Sin embargo, es necesario excluir muchos de estos para obtener información que ayude con el desarrollo del proyecto. Por este motivo se definen criterios de inclusión y exclusión para seleccionar los artículos más relevantes dentro de esta investigación. Los criterios de inclusión y exclusión se presentan en la **Tabla 2**.

Inclusión	Exclusión
<ul style="list-style-type: none"> • Usan LMC para adquirir datos • Usan KNN para procesar los gestos capturados con LMC • Controlan el avatar con LMC • Implementan un ambiente virtual en 3D 	<ul style="list-style-type: none"> • No controlan avatares con LMC • No usan algoritmos de Machine Learning para el reconocimiento de gestos • No manipulan avatares para interactuar con el VW • Artículos Repetidos • No son implementaciones • Artículos que son revisiones de literatura

Tabla 2. Criterios de Inclusión y Exclusión

Elaborado por: El Investigador

El objetivo principal de esta investigación es definir si existe una plataforma de interacción virtual donde se controle un avatar por medio de gestos de la mano capturados por LMC y procesados con KNN. Es por esto por lo que se excluyen 12.407 artículos porque si bien es cierto que usan algoritmos de Machine Learning, no usan un avatar para interactuar con el VW. Como resultado, se obtienen 477 artículos a los cuales se aplican los criterios de exclusión para obtener 7 artículos, de los cuales, únicamente 4 realizan implementaciones.

A pesar de que no se han encontrado trabajos que cumplan exactamente con el objetivo principal de este proyecto, se detallan a continuación los trabajos relacionados más relevantes:

En [8] proponen una aplicación del museo virtual denominada E-Temple, donde se presentan herramientas de adoración hindú con la experiencia real del templo. Incluyen características como guías virtuales, simulaciones, entornos e interacciones 3D, creando una interacción más inmersiva. Donde los usuarios pueden explorar el museo virtual mediante el uso de gestos de las manos, capturados por el sensor LMC. Todas las colecciones creadas en 3D fueron reconstruidas sobre objetos reales mediante la técnica fotogramétrica. En este sentido, utilizan el clasificador K-Nearest Neighbor (KNN) con distancia Euclidiana debido a que los cálculos no son tan complejos y, además son eficientes para ser aplicados en grandes cantidades de datos. Para el desarrollo de la aplicación utilizan 3 procesos: El primero es la adquisición de datos, donde repiten cada gesto 10 veces para cada patrón de la mano. El valor tomado en este primer proceso está en función de la posición de los dedos y de la palma de la mano. Los gestos que registran son: Tap Gesture, utilizado para seleccionar y activar objetos dentro de la aplicación. Change Volume, utilizado para ajustar el volumen del sonido de fondo dentro de la aplicación. Ring Gesture, utilizado para desactivar la guía virtual y que los usuarios puedan explorar el museo virtual libremente. Wave to Activate, utilizado para activar la guía virtual dentro de la aplicación. El segundo proceso es la extracción de características, donde toman valores en el ángulo entre los dedos, el ángulo entre las puntas de los dedos y el centro de la palma, la distancia entre las puntas de los dedos y el centro de la palma, y la evaluación entre las puntas de los dedos y el centro de la palma. El resultado de estos valores se utiliza en los cálculos de la clasificación. Además, utilizan 4 tipos de características: El primero es el ángulo entre los dedos. El segundo es el ángulo entre las yemas de los dedos y las palmas de las manos. El tercero es la distancia entre las yemas de los dedos y las palmas. El cuarto es la elevación entre las yemas de los dedos y las palmas. El tercer proceso es la clasificación: se hacen predicciones para calcular la similitud entre dos datos disponibles, puesto que es necesario encontrar las muestras de datos más similares entre los datos de entrenamiento. Se realizaron 5 pasos para el algoritmo KNN: El primero paso es determinar el valor de vecinos más cercanos (K). El segundo paso es calcular el cuadrado de la distancia euclidiana de un objeto contra los datos de entrenamiento. El tercer paso es ordenar los resultados en orden ascendente. El cuarto es recoger la clasificación del vecino más cercano (Y) basada en el valor de K. El quinto es el uso de Y mayoritario para predecir el resultado. En esta aplicación

utilizaron el IDE Visual Studio 2015 con el lenguaje de programación C# para crear un programa y utilizarlo en la lectura de datos en la adquisición de datos. El cual consiste en 9 columnas: La primera es el recuento del tiempo por movimiento de cuadro capturado en unidades de jj:mm:dd.fff donde fff es milisegundos. La segunda es el ID del cuadro capturado por el LMC, en unidades de 100 cuadros por segundo. La tercera es el número de dedos capturados por el LMC. La cuarta es #0 que representa el valor del pulgar. #1 representa el dedo índice. #2 representa el dedo medio. #3 representa el dedo anular. #4 representa el dedo meñique. Donde en cada uno se encuentran valores vectoriales tridimensionales correspondientes a los ejes x, y, y z. La novena columna es la posición de la palma de la mano, que contiene la posición de los vectores tridimensionales desde el centro de la palma. Realizaron 7 procesos al dataset. El primero consiste en mezclar todos los datos del dataset. El segundo es definir el 20% del dataset para pruebas. El tercero es crear un dataset con el 80% del dataset total para entrenamiento y el 20% para pruebas. El cuarto es inicializar el dataset de entrenamiento y pruebas como diccionarios de datos. El quinto es llenar el diccionario de datos de entrenamiento con el dataset de entrenamiento. El sexto es llenar el diccionario de datos de prueba con el dataset de prueba. El séptimo es pasar los datos del dataset de prueba uno por uno para hacer la predicción. Finalmente, realizaron 5 experimentos con un usuario tanto para entrenamiento como para pruebas, obteniendo una precisión de 99.91%, 98.93%, 99.89%, 100%, y 99.85% respectivamente luego de realizar 5 repeticiones en cada experimento. Por lo tanto, obtienen una precisión promedio de 99.72%.

En [9], proponen una plataforma de realidad virtual basada en la nube, denominada SIGVerse. Donde se propone, además, una arquitectura de software para un simulador de interacción humano-robot (HRI) en entornos de realidad virtual (VR). Esta plataforma está basada en la arquitectura cliente-servidor. Es decir, los usuarios pueden iniciar sesión con un avatar en un mismo servidor desde computadores separadas, e interactuar con robots virtuales. Al estar basado en la tecnología de red integrada de Unity, los clientes y el servidor comparten la misma escena compuesta por modelos de objetos 3D como avatares, robots y muebles. Es posible sincronizar los eventos de cada escena, gracias a la información de los objetos registrados a través de internet. El usuario puede iniciar sesión a través de interfaces de VR como gafas de

realidad virtual, dispositivos de captura de movimientos y auriculares. Dependientemente de la entrada de los dispositivos VR, el comportamiento del usuario se refleja en el avatar mediante scripts de Unity. Por lo tanto, el usuario puede interactuar con el entorno virtual como si fuera un entorno real. Para la aplicación VR, adoptan un middleware de Unity, dado que no es necesario escribir controladores de dispositivos para las interfaces de VR. Además, desarrollan un mecanismo de conexión para vincular Unity y el entorno ROS. Este mecanismo es desarrollado para controlar un robot en un entorno VR, y para que esto funcione, se deben pasar la retroalimentación sensorial y los comandos del robot entre los scripts de Unity y los nodos de ROS; al usar estos recursos basados en ROS, se pueden reutilizar los recursos existentes de robots reales en robots virtuales. Es importante saber que, al enviar grandes cantidades de información desde el dispositivo de entrada, como imágenes de cámaras desde Unity, se crea un cuello de botella en la velocidad de transferencia de datos. Es por esto por lo que proponen una nueva técnica para transferir información en tiempo real basada en formato BSON a través de una conexión TCP/IP. Además, compararon los comportamientos interactivos de robots y avatares en un entorno de realidad virtual y un entorno real, y los resultados de la comparación confirman que los comportamientos de VR de robots y avatares eran casi los mismos que el comportamiento de un robot en un entorno real.

En [10], desarrollan un sistema de reconocimiento de gestos de la mano (HGR) estático que se puede usar en un entorno virtual, donde los usuarios pueden interactuar usando únicamente gestos. El sistema incorpora un algoritmo de aprendizaje automático, debido a que este debe ser suficientemente flexible para reconocer cualquier gesto estático. El entorno de VR se muestra en el Oculus Rift DK2, y los datos de la mano son capturados por el LMC, el cual está montado en el Oculus Rift. Además, incorporan un vector de características de peso ligero de 5 distancias normalizadas de punta a palma y un clasificador k-vecino más cercano (KNN) para detectar el gesto que se está realizando en ese momento. Los datos recibidos se procesan primero en un vector de características ligero y luego son clasificados en un gesto. Para la extracción de características utilizaron un conjunto de 5 distancias normalizadas desde las puntas de los dedos hasta el centro de la palma (TPP), donde estas distancias fueron normalizadas dividiendo cada distancia por la distancia TPP máxima del dedo medio,

restringiendo los valores entre 0 y 1. En cuanto al vector de características, contiene 5 distancias TPP normalizadas que corresponden a cada dedo, donde en la normalización se divide la distancia TPP de cada dedo por la distancia máxima del mismo dedo, en lugar del dedo medio. El algoritmo KNN lo usan por su baja complejidad para clasificar gestos, su alta precisión y su velocidad con un conjunto de datos pequeño. Establecieron un valor bajo de $k=3$ debido al pequeño conjunto de datos y la simplicidad de los gestos en la aplicación. Además, KNN no desecha una entrada por no asemejarse al conjunto de datos, pero siempre devuelve una entrada clasificada. Sin embargo, esto no es del todo correcto, pues se deberían rechazar por completo las entradas que no estén dentro del conjunto de datos. Es por esto por lo que en VR no es apropiado que la aplicación responda a un gesto clasificado incorrectamente cuando debería ser clasificado como desconocido. En este contexto, adaptaron KNN para rechazar los gestos de entrada cuando sean demasiado diferentes de los gestos de entrenamiento. Lo implementaron a través de una función umbral simple, donde si la distancia de Manhattan a los k -vecinos más cercanos supera el umbral, el cual es encontrado empíricamente, entonces el gesto se clasifica como desconocido y la aplicación no responde al mismo. Con el fin de que los gestos persistan entre sesiones, almacenan un archivo de conjunto de datos en el disco. Donde cada vez que la aplicación se inicia, el archivo se lee en memoria de trabajo para ser usado por el algoritmo KNN. En el caso de presentarse un nuevo tipo de gesto, el conjunto de datos se puede volver a escribir en el archivo. Al final, el sistema fue evaluado según varias métricas como el tiempo de respuesta, precisión y usabilidad, utilizando una aplicación de visualización de datos estelares VR de estudio creado en Unreal Engine 4; logrando un tiempo de clasificación de gestos promedio de 0.057 ms con una precisión de 82.5% en 4 gestos distintos (Fist, Point, OpenHand and OK gesture).

En [3], proponen un sistema de ayuda a estudiantes de medicina para aprender las técnicas y habilidades requeridas en la unidad de cuidados intensivos médicos (UCI), brindándoles una experiencia virtual a través de la interacción con los equipos en el ambiente virtual. La interacción virtual con la habitación del hospital se basa en la utilización del Samsung Gear VR y en la implementación del Leap Motion Controller (LMC) que se utiliza para rastrear el movimiento de la mano. Además, utilizan Unity 3D como el motor de desarrollo para construir esta aplicación. En este trabajo,

presentan 3 módulos: 1) creación del entorno virtual. – en este módulo fueron creados objetos en 3D como pacientes, sillas, camas de hospital, camas de UCI, monitores, ventiladores, máquinas de anestesia, etc., integrando el modelo Fbx en Unity 3D. De igual manera, integran la aplicación en Android y se entrega en Samsung Gear VR. Además, la vista 360° de la UCI se experimenta en los auriculares de VR. 2) integración del controlador Gear VR. – el controlador hace que los auriculares sean más ventajosos de utilizar, pues reemplaza todas las actividades que se pueden hacer en el auricular. Este controlador se renderiza en la escena de la UCI virtual. 3) aplicar las capacidades del LMC en la aplicación. – el LMC se puede conectar directamente a la PC o a los auriculares de VR. Este produce alrededor de 200 fotogramas de información reflejada por cada segundo, luego se envía a través de un enlace USB a la PC host. Una de las capacidades del LMC es el gesto Pinch/Grab y el movimiento de objetos. En este contexto, cada máquina de UCI puede moverse hacia el paciente y este realiza todas las actividades con las manos desnudas. Otra capacidad es el reconocimiento de la posición del dedo, aunque esté fuera de la vista del sensor o que esté oculto por objetos externos. Esta aplicación proporciona a los usuarios una experiencia de entrenamiento real en cualquier momento y cualquier lugar.

1.2.3. Fundamentación Teórica

Interacción Hombre-Máquina

La interacción hombre-máquina es la comunicación entre un usuario y una máquina (computador o cualquier dispositivo tecnológico) donde el usuario es el encargado de realizar una serie de comandos que son transmitidos a la máquina a través de periféricos de entrada como teclado, ratón o pantalla táctil. Es la máquina la que interpreta dichos comandos, los procesa y emite una respuesta al usuario a través de periféricos de salida como la pantalla o parlantes. De esta forma, la comunicación entre un hombre y una máquina se vuelve una comunicación bidireccional, porque ambos actuadores envían y reciben información [11].

Inteligencia Artificial (IA)

La IA puede ser definida como el desarrollo de agentes inteligentes, que simulan el comportamiento de los seres humanos. Estos agentes pueden desarrollar rasgos como el conocimiento, aprendizaje e incluso la capacidad de manipular objetos, moverse o resolver problemas [12]. Además, la IA comprende los siguientes aspectos:

Aspectos	Técnicas
<ul style="list-style-type: none">• Conseguir que un sistema tenga razonamiento	<ul style="list-style-type: none">• Razonamiento automatizado• Planificación de pruebas• Resolución de problemas basado en casos
<ul style="list-style-type: none">• Conseguir que un programa aprenda y haga predicciones	<ul style="list-style-type: none">• Aprendizaje automático• Minería de datos• Descubrimiento de conocimiento científico
<ul style="list-style-type: none">• Conseguir que un programa se comunique con seres humanos	<ul style="list-style-type: none">• Procesamiento del lenguaje natural (NLP).
<ul style="list-style-type: none">• Permitir que las máquinas naveguen de forma independiente en un entorno desconocido	<ul style="list-style-type: none">• Planificación• Visión

Tabla 3. Aspectos y Técnicas Comprendidas por la IA

Elaborado por: [12]

Algoritmos de Machine Learning

Los algoritmos de Machine Learning extraen y procesan información de grandes conjuntos de datos, con el objetivo de aprender sin ser programados para realizar tareas específicas. Es decir, adquieren inteligencia a través de la capacidad de tomar decisiones basadas en la naturaleza del aprendizaje o la retroalimentación. Sin embargo, existen problemas ocasionados por el enorme tamaño de la muestra o la gran cantidad de variables, y más cuando se habla de campos como la visión por

computadora, procesamiento de imágenes, minería de datos y análisis estadístico. Es así como, el uso de modelos estadísticos y métodos computacionales es inevitable para la resolución de estos problemas. En conclusión, los algoritmos de Machine Learning imitan la capacidad que tienen los seres humanos de aprender, a través de la experiencia, y cumpliendo una tarea asignada con una mínima o nula ayuda externa [13][14].

KNN (k-Nearest Neighbors)

El Clasificador del Vecino más Cercano (kNN) es un método no paramétrico de aprendizaje supervisado. Tiene características de una estructura simple y una alta precisión, logrando una eficiencia cercana a los mejores clasificadores. kNN memoriza todos los casos posibles del entrenamiento, haciendo que el costo computacional del entrenamiento sea cero. A diferencia del entrenamiento, la predicción sí tiene un alto costo computacional. [15] Además, kNN usa una función de distancia para comparar y obtener la similitud que existe entre la señal desconocida y las señales del conjunto de datos de entrenamiento. Es usado normalmente en el reconocimiento de patrones, clasificación de texto e incluso detección de radicalización [16], [17].

Gestos estáticos

Este tipo de gestos de la mano ocurre durante un momento dado donde la mano tiene una posición bien definida. No se tiene en cuenta los cambios de la posición de la mano en el tiempo que se realiza el gesto y por lo general son definidos previamente por un tipo de modelo guía [7][4][18].

Gestos dinámicos

Los gestos dinámicos se realizan de forma libre, es decir, no siguen ningún tipo de guía. Estos gestos se pueden realizar durante lapsos de tiempo, donde la variación de la posición de la mano es un factor influyente para determinar el tipo de gesto realizado por el usuario [7][4][18].

Entorno Virtual

Un entorno virtual es una simulación del mundo real, diseñado por computadora donde múltiples usuarios pueden acceder a través de sus avatares. Estos entornos forman un espacio para juegos colaborativos, aprendizaje y el trabajo en equipo [19].

Avatar

Cuando se habla de entornos virtuales como los videojuegos, los avatares forman parte fundamental en la relación entre el usuario y el entorno virtual. Estos avatares dependen del contexto que los usuarios prefieran, ya sea una representación realista, abstracta, cómica o autorrepresentación del usuario. Este tiene la finalidad de representar las acciones que el usuario pretende hacer dentro del entorno virtual donde el avatar interactúa con los objetos, teniendo así una percepción similar a la realidad [20].

1.3. Objetivos

1.3.1. Objetivo General

Desarrollar una plataforma de interacción virtual para la navegación en los predios de la UTA del campus Huachi usando un avatar y algoritmos de Machine Learning para mejorar la experiencia de los nuevos estudiantes al conocer lugares de interés.

1.3.2. Objetivos Específicos

- Diseñar una plataforma de interacción virtual que permita reconocer gestos estáticos de la mano para la manipulación del avatar y su interacción con el ambiente a fin de mejorar la experiencia de los nuevos estudiantes de la UTA.
- Diseñar un avatar que se adapte al modelo de reconocimiento de gestos de la mano para optimizar el tiempo de respuesta en los movimientos del avatar.
- Implementar una plataforma de interacción virtual para la navegación en los predios de la UTA del campus Huachi usando un avatar y algoritmos de Machine Learning.

CAPÍTULO II

METODOLOGÍA

2.1. Materiales

2.1.1. Humanos

- Docente Tutor.
- Investigador.
- Estudiantes de la Universidad Técnica de Ambato.

2.1.2. Institucionales

- Repositorios virtuales.
- Biblioteca virtual.

2.1.3. Otros

El proyecto será financiado por el investigador.

No	Detalle	Unidad	Cant.	Valor Unitario	Valor Total
1	Suministros de Oficina	c/u	1	\$5.00	\$5.00
2	Laptop	c/u	1	\$942.857	\$942.857
3	Acceso a Internet	mes	6	\$28.00	\$168.00
4	Leap Motion Controller	c/u	1	\$85.00	\$85.00
5	Visual Studio 2019	c/u	1	\$0.00	\$0.00
6	Matlab R2018b	c/u	1	\$0.00	\$0.00
7	Unity	c/u	1	\$0.00	\$0.00
				Sub Total	\$1200.857
				Imprevistos	\$120.0857
				Total	\$1320.9427

Tabla 4. Recursos Económicos

Elaborado por: El Investigador

2.2. Métodos

2.2.1. Modalidad de la Investigación

Investigación de campo. – Porque se necesita la intervención de dos grupos de estudiantes. Un primer grupo que no conoce presencialmente la UTA y un segundo grupo que sí conoce los predios de la UTA, para comprobar la usabilidad de la plataforma de interacción virtual.

Investigación bibliográfica. – Porque se utilizan fuentes como libros, documentos, artículos científicos, revistas, etc., para la construcción del marco teórico tanto de la variable independiente: Algoritmos de Machine Learning. En la variable dependiente: Experiencia de los nuevos estudiantes al conocer lugares de interés.

Investigación experimental. – Porque se manipulan variables para observar los cambios que se producen en la exactitud con la que el sistema reconoce los gestos de la mano en tiempo real.

2.2.2. Población y muestra

En esta investigación la población son los estudiantes de la UTA, una población finita. Sin embargo, debido a que el modelo de reconocimiento de gestos de la mano debe ser capaz de reconocer el gesto realizado por cualquier persona, y a la complejidad para obtener información del total de estudiantes, se ha decidido trabajar con una muestra para una población desconocida. Para esto, se aplica la siguiente fórmula para calcular la muestra de una población desconocida:

$$n = \frac{Z^2 PQ}{e^2}$$

En donde:

n = tamaño de la muestra

Z = valor relativo determinado por el nivel de confianza

P = probabilidad de éxito

Q = probabilidad de fracaso

e = error estándar

$Z = 1.96$ (Nivel de confianza del 95%)

$p = 0.5$

$q = 0.5$

$e = 0.007$

$n = 196$

Luego de realizar el cálculo de la fórmula con los valores establecidos por el investigador, se concluye que el tamaño de la muestra es de 196 personas.

2.2.3. Recolección de la información

La recolección de datos para el presente proyecto se dividió en dos fases. La primera se basa en la recolección de información para el modelo de reconocimiento de gestos de la mano. El mismo que será usado para la segunda fase. La cual consiste en crear una plataforma de interacción virtual, permitiendo al usuario navegar por los predios de la Universidad Técnica de Ambato.

Para la primera fase, se utilizó un programa desarrollado en Matlab enfocado en la recolección de información. Donde cada usuario realizó 5 gestos con la mano sobre el sensor LMC. El usuario repite cada gesto 30 veces. Los gestos que se reconocen son definidos como OPEN HAND, CLOSE HAND, WAVE IN, WAVE OUT y PINCH. En cada repetición se almacenan la velocidad, dirección y la posición espacial en el eje x, y, z, por cada dedo. De esta forma se crea un dataset que se utiliza para entrenar y probar el modelo de Machine Learning [18].

La segunda fase se basa en la recolección de datos por medio de encuestas dirigidas a personas que conocen y que no conocen presencialmente los predios de la UTA.

Preguntas básicas	Explicación
¿Para qué?	Para cumplir con los objetivos de la investigación
¿De qué personas u objetos?	Estudiantes y personas externas a la UTA.
¿Sobre qué aspectos?	Problemas de personas que nunca han conocido presencialmente los predios de la UTA
¿Quién, Quienes?	Investigador
¿Cuándo?	Abril – septiembre 2023
¿Dónde?	UTA, campus Huachi
¿Cuántas veces?	Una
¿Qué técnicas de Recolección?	Encuestas
¿Con qué?	Cuestionarios

Tabla 5. Técnica de Recolección de Información

Elaborado por: El Investigador

CAPÍTULO III: RESULTADOS Y DISCUSIÓN

3.1. Desarrollo de la propuesta

3.1.1. Diseño de una plataforma de interacción virtual que permita reconocer gestos estáticos de la mano

El diseño de la plataforma de interacción virtual consta de 3 fases. La primera fase consiste en la extracción de un modelo georreferenciado particular de la UTA. En la segunda fase se establecen los edificios con mayor relevancia para los estudiantes, convirtiéndolos en objetos con los cuales el avatar puede interactuar. En la tercera fase, se integra el modelo de Machine Learning utilizado para el reconocimiento de gestos de la mano.

3.1.1.1. Extracción de un modelo georreferenciado (Infraworks).

Se utiliza el programa Infraworks para hacer georreferenciación, es decir, ingresar las coordenadas de un lugar y obtener un diseño 3D del mismo incluyendo edificios, vegetación y todo tipo de construcciones realizadas por el hombre. Aunque este diseño no es tan detallado, sirve como base para que las dimensiones y escalas de los edificios del campus tengan mucho más sentido, y de esta manera lograr una simulación que se asemeje más a la realidad.

3.1.1.2. Transformación de edificios en objetos 3D (Sketchup).

Una vez obtenido el diseño general en 3D del campus, Sketchup es utilizado con el fin de transformar los edificios con mayor relevancia para los estudiantes, en objetos 3D con los cuales el avatar puede interactuar. Esto se lo realiza debido a que el diseño 3D generado por Infraworks solamente forma los edificios como bloques, mas no como objetos con los que el usuario pueda interactuar. Es necesario que el avatar sea capaz de ingresar a estos edificios y tener una mejor perspectiva de cómo son los edificios en la realidad.

3.1.1.3. Integración del Modelo de reconocimiento de gestos de la mano

El modelo de Machine Learning utilizado para este proyecto consiste en 5 módulos (**Figura 1**). El primer módulo es la adquisición de datos, donde se utiliza el sensor

LMC para capturar los gestos realizados por el usuario. El segundo módulo es el pre-procesamiento de los datos adquiridos en la etapa 1. El tercer módulo es la extracción de características. El cuarto módulo es la clasificación, donde se utiliza el clasificador KNN. El último módulo es el post-procesamiento, donde se obtiene una etiqueta predicha dependiendo el gesto realizado por el usuario.

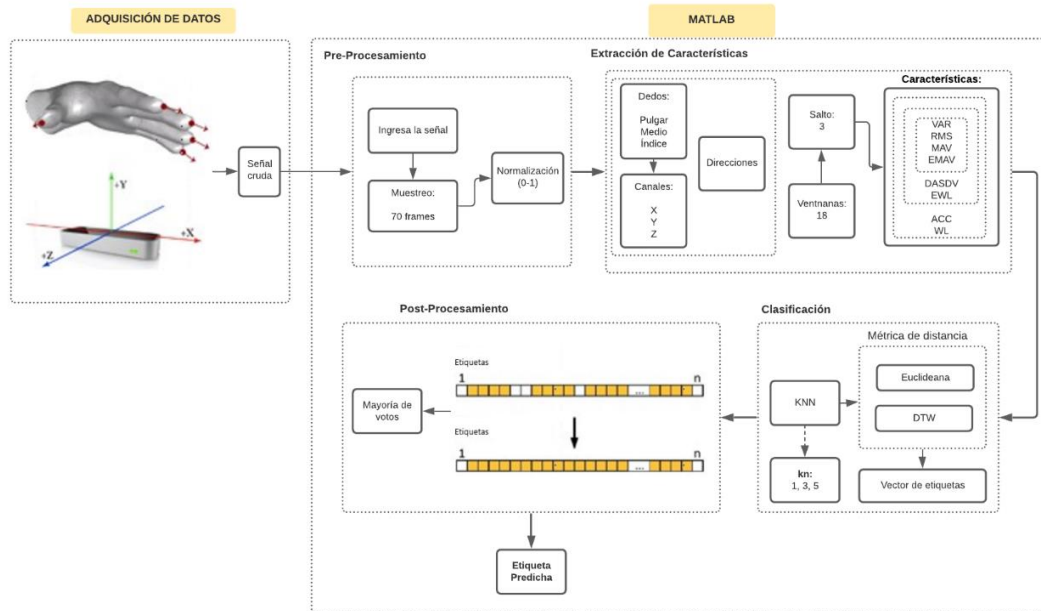


Figura 1. Esquema del modelo de reconocimiento de gestos de la mano

Elaborado por: El Investigador

3.1.1.3.1. Adquisición de datos.

Se utilizó el sensor Leap Motion Controller para capturar los gestos de la mano realizados por el usuario. Cada gesto fue repetido 30 veces. En cada iteración se capturan y almacenan velocidades, direcciones y posiciones en los ejes x, y, z. De esta forma se crea una estructura que será procesada en el segundo módulo.

3.1.1.3.2. Pre-Procesamiento.

Se basa en la utilización de la estructura creada en el módulo de adquisición de datos. Esta estructura es procesada en el caso de que el número de muestras sea menor o mayor que el número de muestras establecido (70 muestras), dando como resultado una señal con el número de muestras estandarizado, y posteriormente se normaliza cada señal.

3.1.1.3.3. Extracción de características.

En este módulo se utilizan funciones de extracción de características. Estas se extraen dentro de cada señal para mejorar la exactitud con la que el modelo reconoce cada gesto.

3.1.1.3.4. Clasificación.

Este módulo consiste en la utilización del clasificador del vecino más cercano (KNN). Este clasificador tiene un costo computacional de entrenamiento igual a cero. Es decir, utiliza un dataset previamente cargado, con el que va comparando cada señal nueva con todas las señales cargadas. Se establecen parámetros usados para mejorar su precisión. Estos parámetros son $kn=1$, $kn=3$, $kn=5$. Siendo $kn=1$ el valor con el cual el clasificador tiene mayor porcentaje de exactitud. Además, utiliza métricas de distancia Euclidiana y DTW para hacer estas comparaciones, obtener la señal con mayor similitud y finalmente crear un vector con todas las etiquetas resultantes.

3.1.1.3.5. Post-Procesamiento.

Este último módulo se basa en recorrer todo el vector de etiquetas obtenido en el módulo de clasificación, para comparar cada una de ellas, y devolver una etiqueta resultante que representa al gesto realizado por el usuario. Finalmente, esta etiqueta es enviada por medio del protocolo de comunicación TCP/IP.

3.1.2. Diseño de una plataforma de interacción virtual que permita reconocer gestos estáticos de la mano para la manipulación del avatar y su interacción con el ambiente

Esta etapa consiste en desarrollar la interacción del avatar con su ambiente. Es decir, programar las acciones que el avatar puede hacer dentro del mundo virtual. En este sentido, se utilizó el lenguaje de programación C#, debido a que es el lenguaje por defecto que se utiliza en aplicaciones desarrolladas en Unity.

En este contexto, el avatar tiene 6 funcionalidades o acciones que desempeñar (**Tabla 6**). La primera acción es caminar hacia adelante y depende del gesto 1 (Open Hand). Además de caminar hacia adelante, el usuario puede detener el movimiento del avatar

(acción 2) haciendo el mismo gesto. Esto quiere decir que el usuario puede dar la instrucción al avatar de caminar o detenerse usando el mismo gesto de la mano. La tercera acción es abrir puertas, para lo cual, el usuario debe realizar el gesto 2 (Open Hand). El usuario debe tener en cuenta que, para dar esta instrucción al avatar, este mismo debe situarse por delante de la puerta que se desea abrir. Existen 3 edificios cuyas puertas cuentan con esta funcionalidad: Coliseo, FISEI, Biblioteca General. La cuarta acción es girar hacia la izquierda, para lo cual, el usuario debe hacer el gesto 3 (Wave In). En esta acción, el avatar no se traslada, sino que rota en su propio eje. La quinta acción es girar hacia la derecha, para lo cual, el usuario debe hacer el gesto 4 (Wave Out). Igual que en la tercera acción, el avatar rota, no se traslada. La última acción es retroceder, para lo cual, el usuario debe hacer el gesto 5 (Pinch). Esta acción no se ejecuta con normalidad, debido a que este gesto tiene un porcentaje muy bajo de exactitud.

Etiqueta	Acción
(1) Open Hand	Caminar hacia adelante, detenerse
(1) Open Hand	Detenerse
(2) Close Hand	Abrir Puertas
(3) Wave In	Girar hacia la izquierda
(4) Wave Out	Girar hacia la derecha
(5) Pinch	Retroceder

Tabla 6. Acciones realizadas por el avatar

Elaborado por: El Investigador

3.1.2.1. Selección del avatar

Los avatares se pueden clasificar en humanoides y no humanoides. Los avatares de tipo humanoide tienen una animación similar a la de un humano. Es decir, el usuario que controla el avatar puede observar todos los gestos que realiza el avatar con cada instrucción. Por el contrario, los avatares de tipo no humanoide, no pueden ser observados por el usuario que los controla. Es decir, el usuario tiene una vista en primera persona del mundo virtual, como si este mismo fuese un avatar interactuando con su entorno. Este proyecto se basa en la interacción del usuario con el avatar y a su vez, la interacción del avatar con su entorno virtual. Es por esto por lo que se decide

utilizar un avatar de tipo humanoide para la interacción del avatar con el mundo virtual.

3.1.2.2. Integración del avatar al mundo virtual (Unity).

Esta última fase consiste en la integración del avatar al mundo virtual. Es decir, diseñar y programar un avatar (**Figura 2**) para que se integre en el mundo virtual. En este sentido, es necesario utilizar animaciones para que el usuario tenga una mejor visualización de las acciones realizadas por el avatar. Estas animaciones son descargadas de Mixamo, sitio web con una amplia colección de personajes y animaciones que puede ser configuradas y descargadas para ser utilizadas en Unity 3D. Las animaciones utilizadas para este proyecto son caminar, girar hacia la izquierda o derecha, y estar de pie.

Es necesario tener en cuenta que para que el avatar funcione correctamente dentro del mundo virtual, debe especificarse como un avatar de tipo Humanoid. Además, debe agregarse un Capsule Collider (**Figura 3**), encargado de detectar las colisiones del avatar frente a los objetos dentro del mundo virtual.



Figura 2. Modelo del avatar

Elaborado por: El Investigador



Figura 3. Capsule Collider aplicado al Avatar

Elaborado por: El Investigador

3.1.3. Implementación de la plataforma de interacción virtual para la navegación en los predios de la UTA del campus Huachi usando un avatar y algoritmos de Machine Learning

Para la implementación de la plataforma, se utilizan los IDE Unity 3D, Matlab, y Visual Studio. Estos se integran entre sí para que la plataforma funcione correctamente (**Figura 4**). En este sentido, Unity utiliza C# como lenguaje de programación para crear las funcionalidades que se ejecutan dentro de su entorno. Estas funcionalidades permiten al avatar interactuar con el ambiente virtual. Las mismas son programadas y se ejecutan en el IDE Visual Studio 2019.

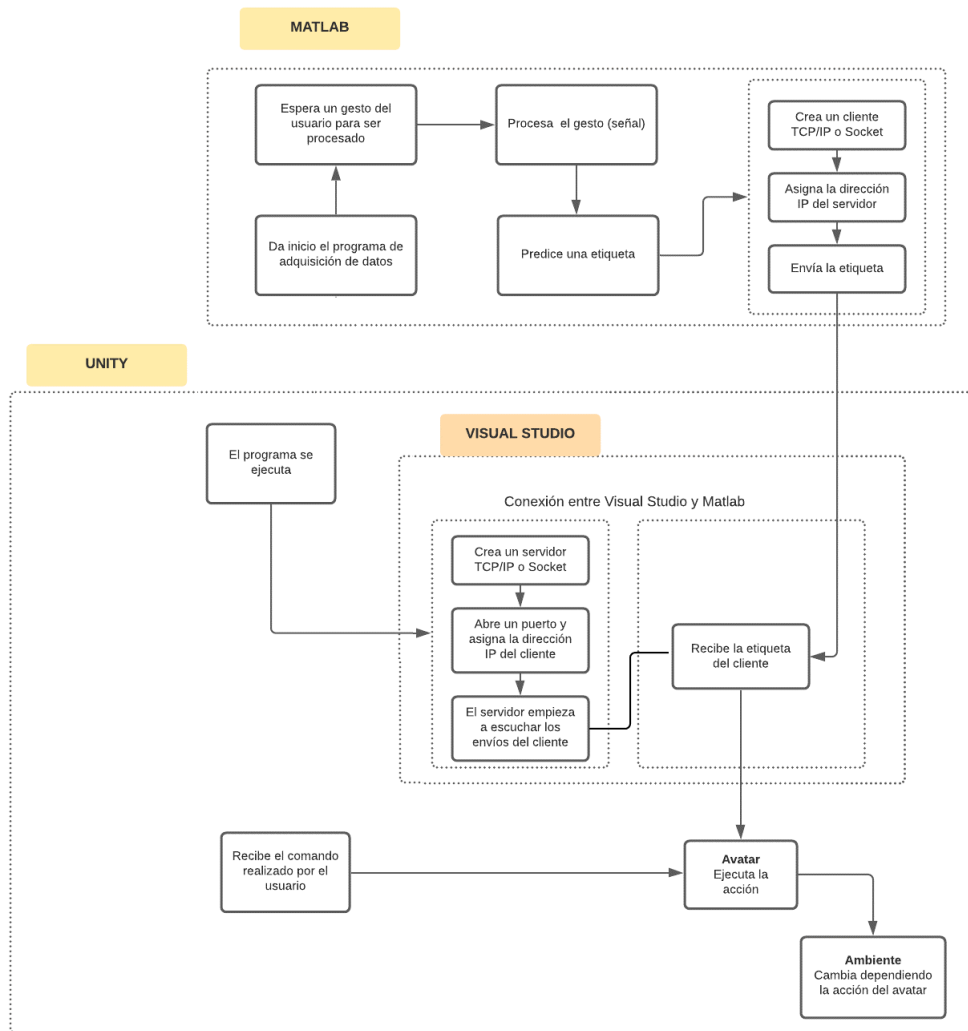


Figura 4. Integración de los IDE

Elaborado por: El Investigador

Por otra parte, para que el avatar pueda moverse necesita de una etiqueta que le diga qué acción se va a ejecutar cada vez. Esta etiqueta es el resultado de una predicción realizada por el modelo de reconocimiento de gestos de la mano, desarrollado en Matlab. Este modelo es el encargado de capturar información, procesarla y emitir una etiqueta resultante que corresponde al gesto realizado por el usuario. Tal como se muestra en la **Figura 1**.

En este contexto, se aplica el algoritmo de Machine Learning conocido como KNN o algoritmo de k vecinos más cercanos, en la fase de clasificación del modelo. Este algoritmo retorna la etiqueta de clase basándose en la mayoría de los votos. Es decir, retorna la etiqueta que se presenta con más frecuencia en el vector de etiquetas. Además, utiliza métricas de distancia Euclidiana y DTW para determinar qué punto

está más cerca del punto de consulta. Tal como se muestra en la **Figura 5**.

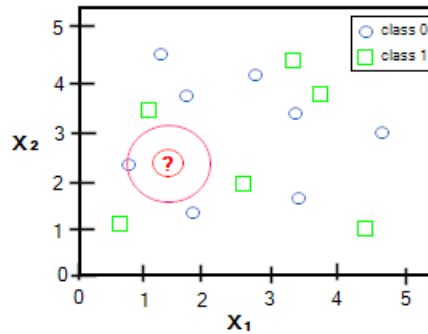


Figura 5. Clasificación de KNN con dos clases (0 y 1)

Elaborado por: El Investigador

Luego de aplicar el algoritmo del vecino más cercano (KNN), se produce la interacción entre Matlab y Visual Studio (C#), mediante el protocolo de comunicación TCP/IP. De esta forma, cada vez que el modelo de reconocimiento predice una etiqueta, esta es enviada desde Matlab hasta Visual Studio. Una vez recibida la etiqueta, el código interpreta que se debe realizar una acción por cada etiqueta recibida. Las acciones se presentan en la **Tabla 6**.

El gesto (5) Pinch no se utiliza para la interacción del avatar debido a que su porcentaje de exactitud es muy bajo (10%) en relación de los otros gestos. Además, el algoritmo interpreta el gesto (5) Pinch como si fuera el gesto (2) Close Hand.

3.2. Experimentación

La experimentación se divide en dos partes. La primera parte consiste en hacer pruebas con el modelo de reconocimiento de gestos de la mano. Mientras que la segunda parte consiste en hacer pruebas con el mundo virtual donde el avatar interactúa con su entorno.

3.2.1. Primera parte

Esta primera parte de la experimentación consiste en probar el modelo de Machine Learning. El mismo que está desarrollado en Matlab, con la versión R2018b. Se ejecuta sobre un computador portátil, cuyas características se detallan en la **Tabla 7**. Como se detalla en la **Figura 1**, este modelo cuenta con 5 módulos. Estos módulos trabajan en

conjunto para predecir una etiqueta que representa el gesto realizado por el usuario.

Características	
Marca	DELL
Sistema Operativo	Windows 10
Procesador	Intel Core i7
Generación	8va
Memoria RAM	8GB

Tabla 7. Características del computador

Elaborado por: El Investigador

De acuerdo con esto, se realizaron 5 pruebas para obtener la combinación de parámetros donde el clasificador KNN obtiene el mayor porcentaje de exactitud. Las pruebas consisten en cambiar el parámetro $kn=5$, $kn=3$, $kn=1$, y realizar 20 repeticiones por cada gesto, realizando un gesto cada vez. Es decir, el usuario hace el gesto y retira la mano del rango de reconocimiento del sensor LMC. De igual manera, se aumenta el número de frames de 200 a 300 y 400.

La primera prueba consiste en asignar a los parámetros los valores de $kn=5$, y $frames=200$. Donde se obtuvo un promedio general de 68% de exactitud realizando 20 repeticiones por cada gesto. Los resultados de esta prueba se muestran en la **Tabla 8**.

GESTO	OPEN HAND (1)	CLOSE HAND (2)	WAVE IN (3)	WAVE OUT (4)	PINCH (5)
ACIERTO	20	15	15	15	3
ERROR	0	5	5	5	17
GESTOS RECONOCIDOS	G.1 = 20 veces	G.2 = 15 veces G.1 = 5 veces	G.3 = 15 veces G.4 = 4 veces G.1 = 1 vez	G.4 = 15 veces G.3 = 4 veces G.1 = 1 vez	G.5 = 3 veces G.2 = 7 veces G.4 = 9 veces G.3 = 1 vez
PORCENTAJE DE EXACTITUD	100%	75%	75%	75%	15%

Tabla 8. Resultados de la primera prueba del modelo

Elaborado por: El Investigador

El número de frames capturados por el sensor LMC influye en la precisión de reconocimiento del modelo, pues se cree que en ocasiones el gesto no es capturado por

completo. Es decir, el número de frames (200), captura solamente una parte del gesto realizado por el usuario. Entonces, se aumenta el número de frames de 300 a 400 y se realiza el mismo principio para capturar datos, procesarlos y evaluarlos con el modelo KNN.

En este contexto, la segunda prueba consiste en asignar a los parámetros los valores de $kn=5$, y subir el número de frames de 200 a 300. Donde se obtuvo un promedio general de 69% de exactitud realizando 20 repeticiones por cada gesto. Los resultados de la segunda prueba se muestran en la **Tabla 9**.

GESTO	OPEN HAND (1)	CLOSE HAND (2)	WAVE IN (3)	WAVE OUT (4)	PINCH (5)
ACIERTO	20	17	16	14	2
ERROR	0	3	4	6	18
GESTOS RECONOCIDOS	G1 = 20 veces	G.2 -> 17 veces G.1 -> 3 veces	G.3 -> 16 veces G.4 -> 3 veces G.1 -> 1 vez	G.4 -> 14 veces G.3 -> 5 veces G.2 -> 1 vez	G.5 -> 3 veces G.2 -> 14 veces G.4 -> 1 veces G.3 -> 2 vez
PORCENTAJE DE EXACTITUD	100%	85%	80%	70%	10%

Tabla 9. Resultados de la segunda prueba del modelo

Elaborado por: El Investigador

La tercera prueba consiste en asignar a los parámetros los valores de $kn=5$, y subir el número de frames de 300 a 400. Donde se obtuvo un promedio general de 71% de exactitud realizando 20 repeticiones por cada gesto. Los resultados de la tercera prueba se muestran en la **Tabla 10**.

GESTO	OPEN HAND (1)	CLOSE HAND (2)	WAVE IN (3)	WAVE OUT (4)	PINCH (5)
ACIERTO	20	17	13	19	2
ERROR	0	3	7	1	18
GESTOS RECONOCIDOS	G1 = 20 veces	G.2 -> 17 veces G.3 -> 2 veces G.1 -> 1 vez	G.3 -> 13 veces G.4 -> 4 veces G.1 -> 3 veces	G.4 -> 19 veces G.3 -> 1 vez	G.5 -> 2 veces G.2 -> 11 veces G.4 -> 1 veces G.3 -> 6 vez

PORCENTAJE DE EXACTITUD	100%	85%	65%	95%	10%
--------------------------------	------	-----	-----	-----	-----

Tabla 10. Resultados de la tercera prueba del modelo

Elaborado por: El Investigador

La cuarta prueba consiste en asignar a los parámetros los valores de frames=400, y bajar el número de kn de 5 a 3. Donde se obtuvo un promedio general de 69% de exactitud realizando 20 repeticiones por cada gesto. Los resultados de la cuarta prueba se muestran en la **Tabla 11**.

GESTO	OPEN HAND (1)	CLOSE HAND (2)	WAVE IN (3)	WAVE OUT (4)	PINCH (5)
ACIERTO	20	18	16	14	1
ERROR	0	2	4	6	19
GESTOS RECONOCIDOS	G1 = 20 veces	G.2 -> 18 veces G.3 -> 1 vez G.4 -> 1 vez	G.3 -> 16 veces G.2 -> 2 veces G.1 -> 2 veces	G.4 -> 14 veces G.3 -> 5 veces G.1 -> 1 vez	G.5 -> 1 vez G.2 -> 18 veces G.3 -> 1 vez
PORCENTAJE DE EXACTITUD	100%	90%	80%	70%	5%

Tabla 11. Resultados de la cuarta prueba del modelo

Elaborado por: El Investigador

La quinta prueba consiste en asignar a los parámetros los valores de frames=400, y bajar el número de kn de 3 a 1. Donde se obtuvo un promedio general de 62% de exactitud realizando 20 repeticiones por cada gesto. Los resultados de la quinta prueba se muestran en la **Tabla 12**.

GESTO	OPEN HAND (1)	CLOSE HAND (2)	WAVE IN (3)	WAVE OUT (4)	PINCH (5)
ACIERTO	20	18	14	8	2
ERROR	0	2	6	12	18
GESTOS RECONOCIDOS	G1 = 20 veces	G.2 -> 18 veces G.1 -> 2 veces	G.3 -> 14 veces G.4 -> 6 veces	G.4 -> 8 veces G.3 -> 12 veces	G.5 -> 2 veces G.2 -> 17 veces G.1 -> 1 vez
PORCENTAJE DE EXACTITUD	100%	90%	70%	40%	10%

EXACTITUD					
-----------	--	--	--	--	--

Tabla 12. Resultados de la quinta prueba del modelo

Elaborado por: El Investigador

El sensor LMC captura las posiciones espaciales, distancia y velocidad en los ejes x, y, z, de cada dedo de la mano que se posiciona sobre la superficie de reconocimiento. Esto quiere decir que mientras el usuario mantiene su mano, el sensor LMC sigue capturando y almacenando grandes cantidades de información que será procesada por el clasificador.

Finalmente, se concluye que la mejor combinación de parámetros para el clasificador es con frames=400 y kn=5. Obteniendo un porcentaje de exactitud del 100% en el gesto 1 (Open Hand), el 85% en el gesto 2 (Close Hand), el 65% en el gesto 3 (Wave In), el 95% en el gesto 4 (Wave Out), y el 10% en el gesto 5 (Pinch), tal como se muestra en la **Tabla 10**. Los gestos 1, 2, 3, y 4, son los que presentan el porcentaje más alto de exactitud. Por lo tanto, estos gestos son utilizados para representar las acciones (**Tabla 4**) que el avatar realiza durante la interacción con su entorno virtual. Además, se obtiene una media de 71% de exactitud en la tercera prueba. La misma combinación se usa para el resto de las pruebas con el mundo virtual.

Una vez terminado el proceso de reconocimiento de gestos de la mano, el modelo se encarga de enviar la etiqueta predicha por medio del protocolo TCP/IP. Para el envío de datos, a través del puerto 55001, se utiliza la dirección IP 127.0.0.1 o el prefijo localhost, que hace referencia a la dirección IP local del dispositivo donde se está ejecutando el modelo.

3.2.2. Segunda parte

La segunda parte de la experimentación consiste en hacer pruebas con el modelo 3D de la UTA. Este modelo 3D fue desarrollado en el IDE Unity, con la versión 2021.3.26f1. Esta versión fue instalada en Unity Hub, aplicación encargada de administrar los proyectos indistintamente de la versión en la que fueron creados. A pesar de ser la aplicación donde el usuario interactúa con el avatar, esta depende de otras dos aplicaciones utilizadas para la creación del entorno virtual. La primera aplicación es Infraworks, versión 2021. Esta aplicación es la encargada de, por medio

de georreferenciación, suministrar un modelo 3D de cualquier lugar existente en el mapa, en este caso, la Universidad Técnica de Ambato. El modelo 3D generado en Infracworks se muestra en la **Figura 6**.

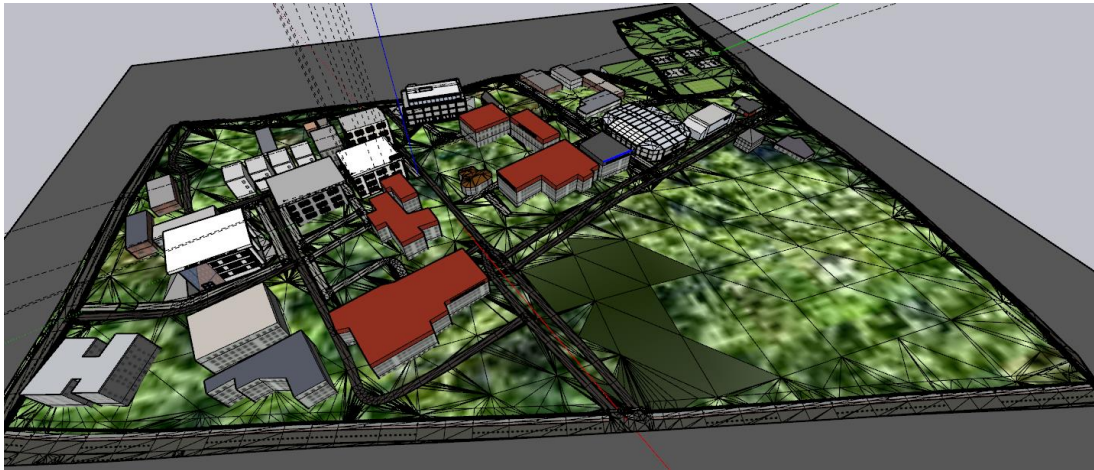


Figura 6. Modelo 3D generado en Infracworks

Elaborado por: El Investigador

Una vez exportado el modelo 3D, se procede a importarlo en la segunda aplicación, Sketchup versión 2022. Sketchup es utilizada para diseñar los edificios de manera que el avatar sea capaz de interactuar con estos y no atravesarlos.

Se realizaron varias pruebas con el mismo ordenador cuyas características se detallan en la **Tabla 7**. Estas pruebas consisten en controlar el avatar y moverlo por el entorno virtual, es decir, las instalaciones de la UTA. Cabe resaltar, que existen lugares donde las edificaciones son pocas o inexistentes, tales como el estadio o la entrada principal de la UTA. Es en estos lugares donde el avatar puede moverse con total normalidad, dando un número mínimo de 60 y un máximo de 130FPS (Fotogramas por Segundo). Cuando se habla de FPS, se refiere a la rapidez con la que se actualizan las imágenes dentro del mundo virtual. Es decir, mientras más alto sea el número de FPS, la experiencia que tiene el usuario controlando el avatar, será más fluida y realista. En cambio, en los lugares donde se encuentran los edificios con más componentes, como la biblioteca, el coliseo, la FISEI y el DIBESAU, el número de FPS no sobrepasa los 7. Esta disminución de FPS, produce un retardo en los movimientos y reacción del avatar, luego de recibir la instrucción (gesto de la mano) del usuario.

En este contexto, se realizó una breve investigación sobre las formas de mejorar la interacción del avatar con su entorno, o de reducir el retardo en la reacción del avatar en relación con la instrucción dada por el usuario. Una de las formas de mejorar la interacción, es reduciendo las elevaciones del modelo 3D. Para esto, se utilizó la funcionalidad Geolocalización de Sketchup. El modelo 3D generado en Sketchup se muestra en la **Figura 7**.



Figura 7. Modelo 3D generado en Sketchup

Elaborado por: El Investigador

Esta funcionalidad, permite ingresar las coordenadas de un lugar y obtener un modelo de referencia. A diferencia de Infracore, esta opción solamente proporciona la base de un lugar, sin elevaciones o edificios. Se utilizó esta funcionalidad para, a partir del modelo de referencia, diseñar el campus Huachi sin elevaciones. Esto disminuye el retardo en la mayoría de los lugares donde existen elevaciones. Sin embargo, el número de FPS solo incrementa de 7 a 9. Sin embargo, la reacción del avatar sigue siendo lenta. Por este motivo se procede a reducir los componentes, dejando los edificios como modelos de referencia.

El recorrido del avatar por las instalaciones del campus Huachi se realiza de forma normal. Sin embargo, el recorrido por la FISEI se limita a la primera planta, teniendo todo el edificio como referencia de la realidad. En el DIBESAU, el avatar se limita a subir por el estacionamiento, utilizando el camino construido para los automóviles,

dejando un pequeño edificio en la parte superior que simula las instalaciones originales. Luego de realizar estos cambios, el número máximo de FPS incrementó de 7 a 11. Esto brinda una mejor experiencia al usuario al momento de dar instrucciones para que el avatar interactúe con su entorno virtual.

3.3. Resultados de las pruebas con la plataforma

La plataforma fue probada por un grupo de 15 personas. Considerando que cada persona realizó un promedio de 30 gestos durante todo el recorrido, se obtiene un total de 450 gestos o instancias en las pruebas realizadas. Además, este grupo de personas está generalizado teniendo en cuenta que la plataforma debe funcionar de igual manera para las personas sin importar su edad. Enfocándose prácticamente en si conocen o no las instalaciones de la UTA.

En este sentido, el grupo de 15 personas se dividió en 2. Un grupo de 5 personas que sí conoce la UTA, y el grupo restante de 10 personas que no conocen la UTA. Se da mayor prioridad a las personas que no conocen la UTA. Teniendo en cuenta que, para tener una mejor evaluación de la experiencia de los usuarios durante el recorrido virtual, es necesario disponer de usuarios que no conozca la UTA.

Como se muestra en la **Figura 8**, las pruebas se dividieron en 3 etapas. La primera etapa consiste en responder 6 preguntas sobre el interés de las personas en utilizar la plataforma. La segunda etapa consiste en probar la plataforma. La tercera etapa consiste en responder las últimas 4 preguntas de la encuesta, para evaluar la experiencia de las personas luego de probar la plataforma. Estas preguntas fueron validadas con el coeficiente del Alfa de Cronbach, obteniendo 0.48 para preguntas con escala de Likert. Para preguntas dicotómicas, se obtiene un valor de 0.36 con el coeficiente de Kuder-Richardson.



Figura 8. Esquema de las pruebas realizadas

Elaborado por: El Investigador

La primera parte de las preguntas de la encuesta se muestra en el **Anexo 1**. En este contexto, la primera pregunta, donde se pide al usuario ingresar su nombre, no se la toma en cuenta debido a que el interés de una persona hacia una aplicación no depende del nombre que esta posea. El resultado de las siguientes 5 preguntas se muestra en la **Tabla 13**. Donde:

P2 = ¿Qué edad tiene?

P3 = ¿Pertenece a la Universidad Técnica de Ambato?

P4 = ¿Conoce las instalaciones del campus Huachi de la Universidad Técnica de Ambato?

P5 = ¿Le interesaría el desarrollo de una plataforma de interacción virtual del campus Huachi de la Universidad Técnica de Ambato?

P6 = ¿Estaría dispuesto a usar una plataforma de interacción virtual, para conocer las instalaciones del campus Huachi de la Universidad Técnica de Ambato?

Encuestados	P2	P3	P4	P5	P6
1	Mayor a 18 años	No	Si	Me interesa mucho	Si
2	Mayor a 18 años	Si	Si	Me interesa mucho	Si
3	Mayor a 18 años	No	Si	Me interesa mucho	Si

4	Mayor a 18 años	No	Si	Me interesa	Si
5	Mayor a 18 años	Si	Si	Me interesa	Si
6	Mayor a 18 años	No	No	Me interesa	Si
7	Mayor a 18 años	No	No	Me interesa	Si
8	Mayor a 18 años	No	No	Me interesa	Si
9	Mayor a 18 años	No	No	Tal vez	Si
10	Menor de 18 años	No	No	Tal vez	Si
11	Mayor a 18 años	No	No	Me interesa mucho	Si
12	Mayor a 18 años	No	No	Me interesa mucho	Si
13	Mayor a 18 años	No	No	Me interesa mucho	Si
14	Mayor a 18 años	No	No	Me interesa	Si
15	Mayor a 18 años	No	No	Tal vez	Si

Tabla 13. Resultados de la primera parte de la encuesta

Elaborado por: El Investigador

En la **Figura 9**, se muestran los resultados de la pregunta 2. Considerando que **1=Menor de 18 años, 2=18 años, 3=Mayor de 18 años**.

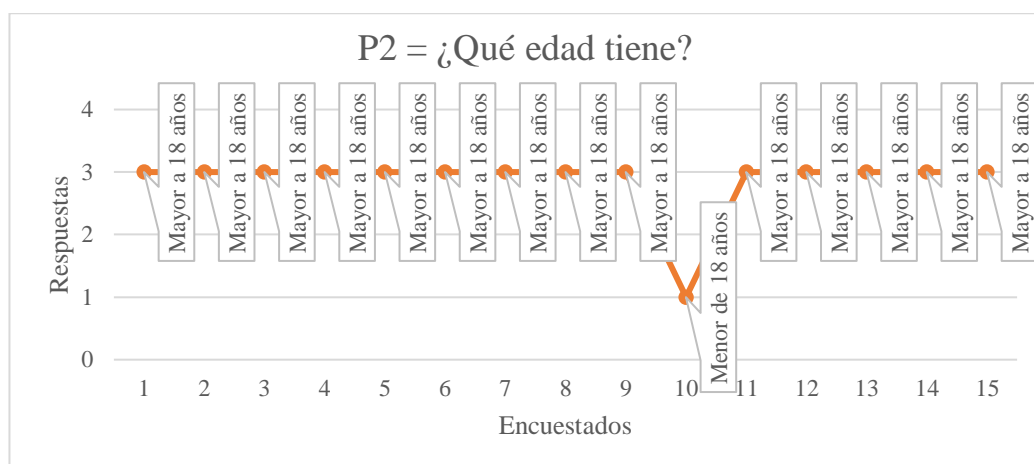


Figura 9. Respuestas - Pregunta 2

Elaborado por: El Investigador

Se obtiene como resultado que 14 encuestados (93.33%) es mayor de 18 años y solamente 1 encuestado (6.67%) es menor de 18 años. Este dato es importante debido al impacto que tiene la edad a la hora de decidir si se interesa o no por el uso de una aplicación que le permita visitar lugares de forma virtual. Además, se considera que por lo general los estudiantes terminan la educación secundaria, antes de cumplir los 18 años. Formando parte del grupo que tendría más interés en utilizar la plataforma para conocer la UTA, debido a su interés en acceder a una institución de educación superior como es la UTA.

En la **Figura 10**, se muestran los resultados de la pregunta 3. Considerando que esta pregunta es de tipo dicotómica. Es decir, solamente tiene dos posibles respuestas, en este caso, **1=Sí, 2=No**.

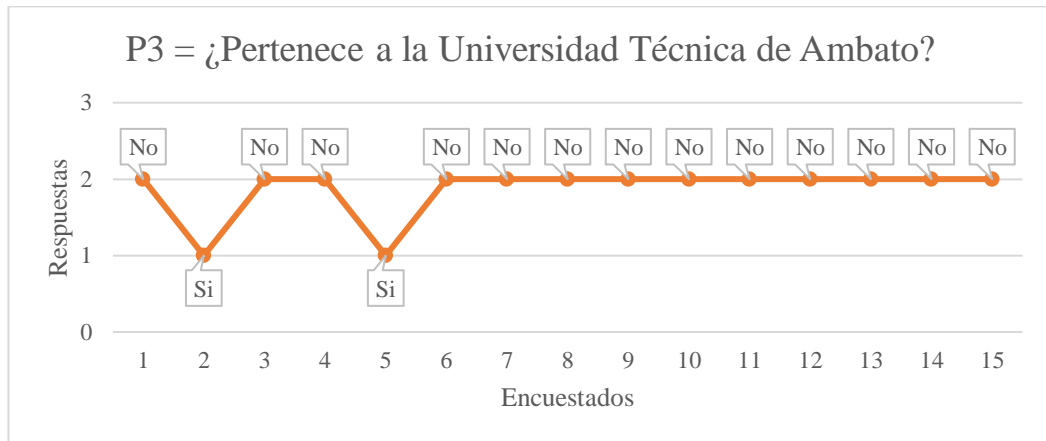


Figura 10. Respuestas - Pregunta 3

Elaborado por: El Investigador

Se obtiene como resultado que 13 encuestados (86.67%) no pertenecen a la UTA. Mientras que 2 personas (13.33%) sí pertenecen a la UTA. Considerando que los primeros 5 encuestados sí conocen la UTA, se puede dar el caso que estas mismas personas no pertenezcan a la UTA. Esto se debe a que en algún momento fueron estudiantes o simplemente acudieron a las instalaciones para acceder a algún servicio que ofrece la UTA. Tales como el Centro de Idiomas, gimnasio, piscina, coliseo, entre otros. En este contexto, las personas que no pertenecen a la UTA, pero que sí la conocen, tienen al menos una referencia de cómo son las instalaciones.

En la **Figura 11**, se muestran los resultados de la pregunta 4. Considerando que esta pregunta dicotómica, donde **1=Sí, 2=No**.

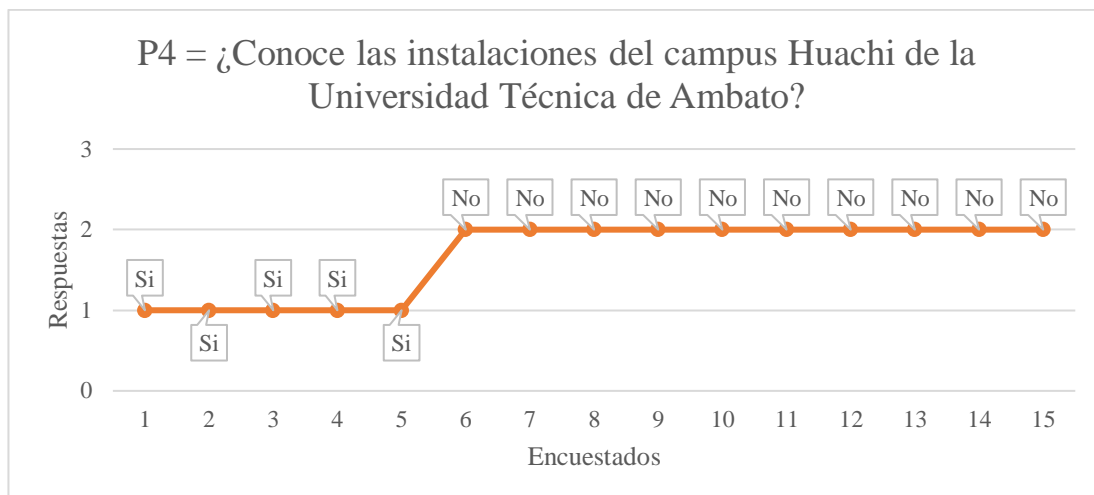


Figura 11. Respuestas - Pregunta 4

Elaborado por: El Investigador

Se obtiene como resultado que 10 encuestados (66.67%) no conoce las instalaciones de la UTA. Y los 5 encuestados restantes (33.33%) sí conoce la UTA. Estas respuestas son válidas y se utilizan como punto de partida para la etapa 2 de la experimentación. Debido a que la población de 15 personas encuestadas se divide en 2 grupos de 5 y 10 personas respectivamente. Donde el primer grupo sí conoce la UTA y el segundo grupo no. En este contexto, los dos grupos tienen que probar la plataforma. Sin embargo, se da prioridad al grupo de personas que no conocen la UTA, porque se evalúa la experiencia que tienen al enfrentarse a un ambiente totalmente desconocido.

En la **Figura 12**, se muestran los resultados de la pregunta 5. Considerando que **1= Me interesa mucho**, **2= Me interesa**, **3= Tal vez**, **4=No me interesa**. Esta pregunta de escala Likert, quiere decir que sus respuestas no se cierran a dos únicas respuestas.

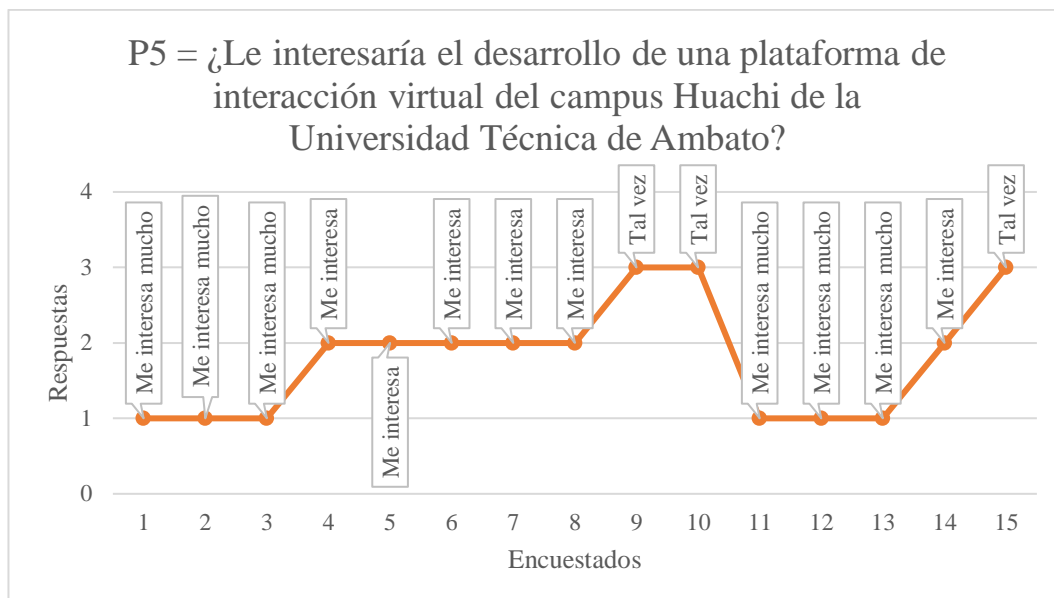


Figura 12. Respuestas - Pregunta 5

Elaborado por: El Investigador

Se obtiene que a 6 encuestados (40.00%) les interesa mucho, a 6 encuestados (40.00%) le interesa, 3 encuestados (20.00%) responde con un Tal vez, y el 0.00% o ninguna persona ha respondido que no le interesa. Considerando que las respuestas 1 y 2 reflejan resultados positivos sobre el interés del encuestado por la plataforma. Se concluye que 12 encuestados sí muestra interés por la plataforma. Dando un resultado positivo para seguir con el desarrollo del proyecto.

En la **Figura 13**, se muestran los resultados de la pregunta 6. Considerando que esta pregunta dicotómica, donde **1=Sí**, **2=No**.

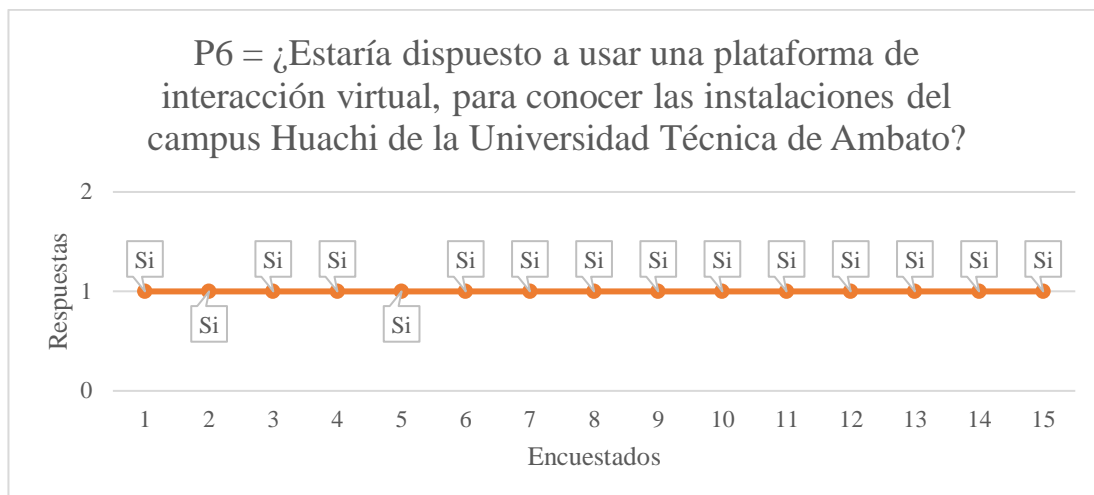


Figura 13. Respuestas - Pregunta 6

Elaborado por: El Investigador

Se obtiene como resultado que el 100% de los encuestados estaría dispuesto a utilizar la plataforma. Este resultado es un reflejo del interés y la disposición que tienen las personas por utilizar plataformas que les permitan conocer lugares de forma virtual, antes de conocerlos físicamente. En este caso, las personas interesadas son personas que conocen y no conocen las instalaciones de la UTA.

Para la segunda etapa de las pruebas, se toman como base las respuestas de la primera parte de la encuesta. Esta etapa consiste en probar la plataforma. Para lo cual se parte del grupo de 15 personas a las cuales se aplicó la encuesta. Este grupo se divide en 2 subgrupos. El primero está conformado por 5 personas que sí conocen la UTA, y el segundo por 10 personas que no conocen la UTA. Para los dos subconjuntos de personas, se les facilitó una breve explicación sobre cuáles son los gestos que el avatar reconoce como instrucciones, y cómo deben ser ejecutados.

El primer subconjunto de 5 personas tuvo 5 minutos iniciales para socializarse con el aplicativo. Luego, el mismo subconjunto tuvo 10 minutos cada uno para hacer un recorrido desde la entrada principal hasta la puerta cercana al edificio de Cultura Física, y posteriormente dirigirse hacia la puerta cercana a la biblioteca general. Este recorrido se basa en que las personas que sí conocen la UTA tienen claro el entorno por donde mover el avatar. Sin embargo, el resultado es que solamente 2 personas concluyen el recorrido, mientras que los 3 restantes tuvieron dificultades para controlar el avatar. El resultado de estas pruebas no generaliza a una población, dado que las

pruebas se ejecutan con 5 personas. Sin embargo, se puede generalizar que los problemas presentados por las 3 personas que no concluyen el recorrido son los mismos problemas que enfrentan las 12 personas restantes usadas para la experimentación. Esto se debe a que a pesar de recibir una explicación de cómo realizar cada gesto, las personas siguen equivocándose.

El segundo subconjunto conformado por 10 personas que no conocen las instalaciones de la UTA tuvo 5 minutos para socializarse con el aplicativo. Terminada esta primera parte, el mismo subconjunto tuvo 20 minutos en los cuales tenía que navegar libremente por las instalaciones. Considerando que, al no conocer el entorno, deben explorar los caminos por los cuales el avatar se puede trasladar. Posteriormente, se les pidió hacer un breve recorrido por las instalaciones originales luego de terminar la prueba con la plataforma. Esto se realiza para tener una mejor perspectiva de lo que refleja la plataforma en relación con el campus original, y para poder responder la segunda parte de la encuesta.

La tercera etapa de las pruebas consiste en evaluar la experiencia de los encuestados luego de utilizar la plataforma. Para esta etapa, los dos subgrupos de personas deben responder las últimas 4 preguntas de la encuesta. Las preguntas se muestran en el **Anexo 2**. El resultado de las 4 preguntas se muestra en la **Tabla 14**. Donde:

P7 = Califique su experiencia utilizando la plataforma de interacción virtual del campus Huachi de la Universidad Técnica de Ambato

P8 = ¿Volvería a usar la plataforma de interacción virtual, para conocer el campus Huachi de la Universidad Técnica de Ambato?

P9 = ¿Recomendaría el uso de la plataforma de interacción virtual, para conocer el campus Huachi de la Universidad Técnica de Ambato?

P10 = ¿Tiene usted alguna sugerencia?

Encuestados	P7	P8	P9	P10
1	Buena	Si	Si	
2	Muy buena	Si	Si	
3	Buena	Si	Si	Ninguna

4	Buena	No	Tal vez	Capacitaciones para saber manejar dicha aplicación
5	Buena	Si	Tal vez	Más velocidad con el avatar para un recorrido completo
6	Muy buena	Si	Si	Que el muñequito salte
7	Muy buena	Si	Si	No
8	Buena	Si	Si	
9	Buena	Si	Si	Que pueda saltar
10	Muy buena	Si	Si	Ninguna
11	Muy buena	Si	Si	Sigan con empeño.
12	Muy buena	Si	Si	Mayor cantidad de gestos para más interacciones en el campus
13	Muy buena	Si	Si	Capacitación para el uso de ello mismo
14	Muy buena	Si	Si	
15	Regular	No	Tal vez	Los movimientos son muy cansados

Tabla 14. Resultados de la segunda parte de la encuesta

Elaborado por: El Investigador

En la **Figura 14**, se muestran los resultados de la pregunta 7. Donde: **1**=Muy buena, **2**=Buena, **3**=Regular.

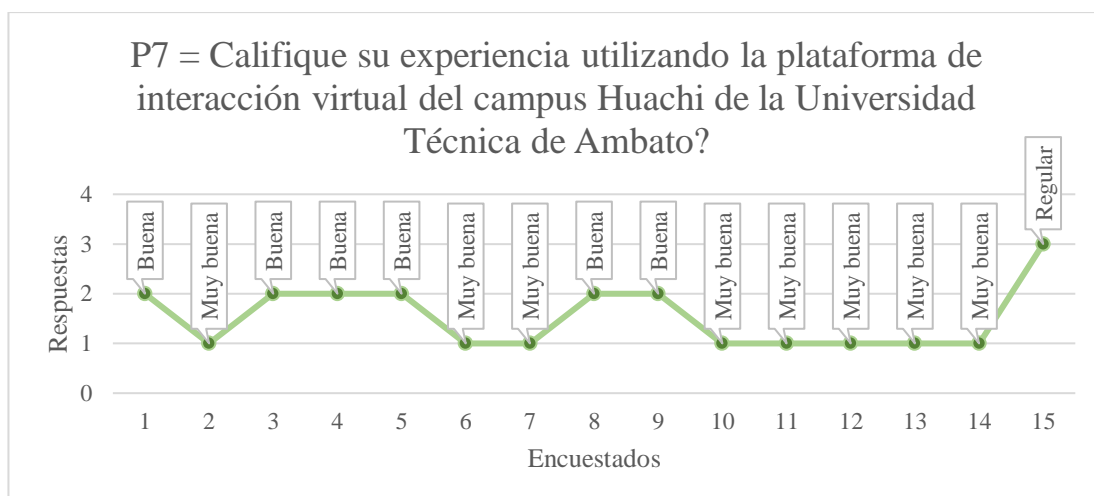


Figura 14. Respuestas - Pregunta 7

Elaborado por: El Investigador

Se obtiene como resultado que 8 encuestados (53.33%) tuvo una experiencia muy buena, 6 encuestados (40.00%) tuvo una experiencia buena, y 1 encuestado (6.67%) tuvo una experiencia regular. Considerando que, del grupo de 10 personas que no conocen la UTA, 5 recorrieron las instalaciones físicamente antes de realizar la segunda parte de la encuesta. Se concluye que, 5 de las 15 personas encuestadas tuvieron una buena experiencia en su primera visita a la UTA. Mientras que una sola persona tuvo una experiencia regular, debido a la dificultad que esta presentó para hacer los gestos y, por consiguiente, controlar el avatar.

En la **Figura 15**, se muestran los resultados de la pregunta 8. Considerando que **1=Sí**, **2=No**.

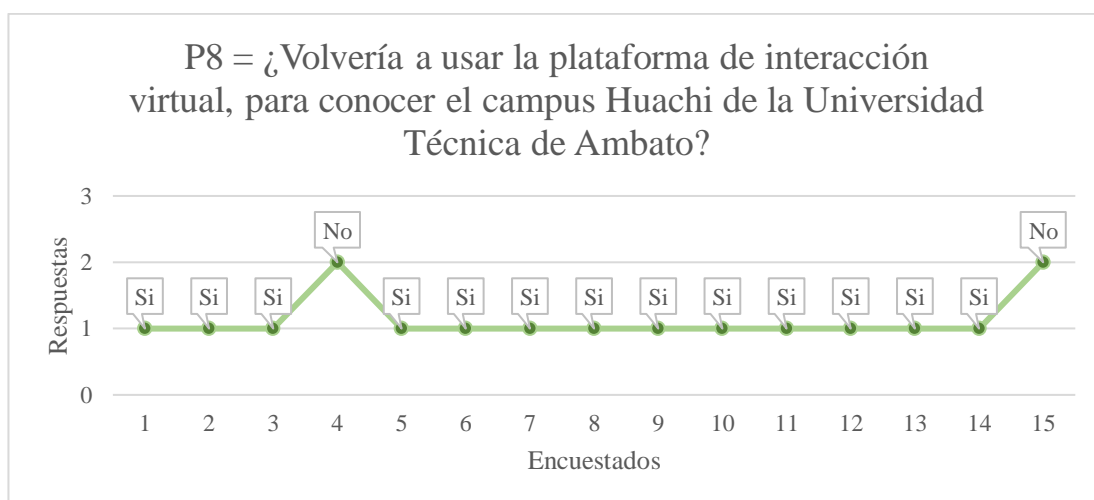


Figura 15. Respuestas - Pregunta 8

Elaborado por: El Investigador

Se obtiene como resultado que 13 encuestados (86.67%) sí volverían a usar la plataforma, y 2 encuestados (13.33%) no la volverían a usar. En este sentido, se tiene que 13 de los 15 encuestados están satisfechos con la plataforma. Y, por lo tanto, la volverían a usar para conocer las instalaciones de la UTA de forma virtual.

En la **Figura 16**, se muestran los resultados de la pregunta 9. Considerando que **1=Sí**, **2=Tal vez**, **3=No**.

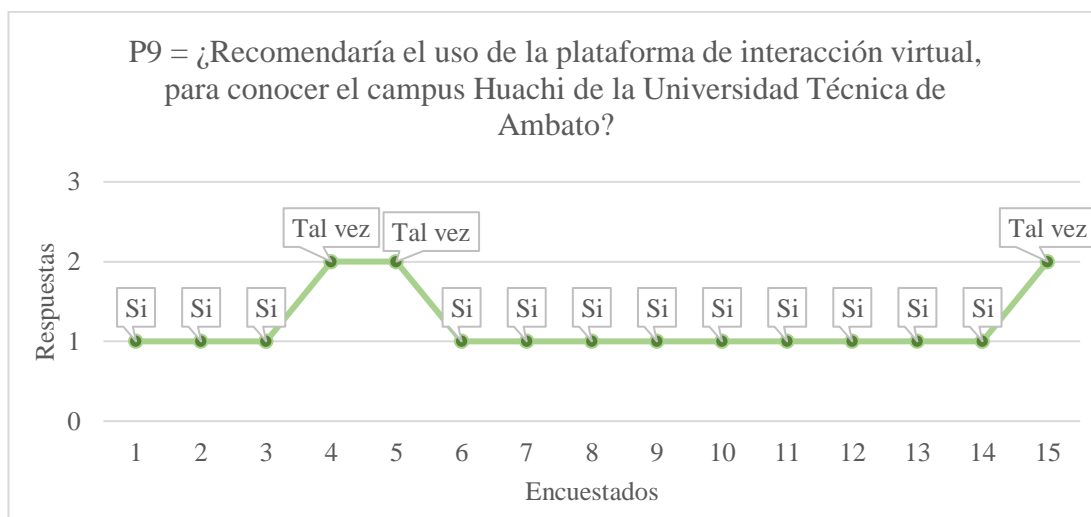


Figura 16. Respuestas - Pregunta 9

Elaborado por: El Investigador

Se obtiene como resultado que 12 encuestados (80.00%) respondieron sí, y 3 encuestados (20%) respondieron tal vez. Considerando que 12 encuestados sí recomendaría el uso de la plataforma, se concluye que la funcionalidad de conocer la UTA de forma virtual es aceptada por al menos 12 de las 15 personas encuestadas.

En la **Tabla 15**, se muestran los resultados de la pregunta 10. Esta pregunta de tipo abierta requiere que el encuestado ingrese una respuesta con sus propias palabras, sin tener ninguna opción que lo limite.

Encuestados	P10 = ¿Tiene usted alguna sugerencia?
1	
2	
3	Ninguna
4	Capacitaciones para saber manejar dicha aplicación
5	Más velocidad con el avatar para un recorrido completo
6	Que el muñequito salte
7	No
8	
9	Que pueda saltar
10	Ninguna
11	Sigan con empeño.

12	Mayor cantidad de gestos para más interacciones en el campus
13	Capacitación para el uso de ello mismo
14	
15	Los movimientos son muy cansados

Tabla 15. Respuestas - Pregunta 10

Elaborado por: El Investigador

Se obtiene como resultado que 7 personas no tienen sugerencias, reflejando la conformidad que tuvieron al momento de utilizar la plataforma. Sin embargo, las 8 personas restantes sí tienen sugerencias. Esto quiere decir que, a pesar de tener una buena experiencia utilizando la plataforma, encuentran problemas que se pueden tratar en el futuro. Destacando entre los problemas más importantes: capacitaciones para manejar la plataforma, aumentar la velocidad en los movimientos del avatar, aumentar el número de gestos de la mano reconocidos por el modelo, aumentar las acciones que pueda realizar el avatar dentro de su entorno virtual. Además de crear un sistema o método que permita al usuario adaptarse a los gestos para hacerlos correctamente y de esa forma, controlar mejor el movimiento del avatar. Existe una última sugerencia que se enfoca en la dificultad que tienen los gestos, provocando cansancio al usuario que utiliza la plataforma. Del mismo modo, se pueden usar estas sugerencias para desarrollar nuevas funcionalidades en futuras versiones.

CAPÍTULO IV

CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones

- Se desarrolló una plataforma de interacción virtual para visitantes primarios. Sin embargo, las pruebas se realizaron con un grupo conformado por 5 personas que sí conocen la UTA, y 10 personas que no la conocen. Esta plataforma funciona con 2 sistemas integrados. El primer sistema utiliza el clasificador KNN para predecir una etiqueta que representa el gesto de la mano, realizado por el usuario. Esta etiqueta se envía por medio del protocolo TCP/IP. El segundo sistema recibe la etiqueta y da la instrucción al avatar para que interactúe con su entorno virtual.
- El avatar utilizado para la interacción con el mundo virtual es uno del tipo humanoide. Este fue programado de tal manera que su reacción a los gestos realizados por el usuario, fuera casi inmediata. Sin embargo, se evidenciaron retardos en la reacción, debido a las limitaciones del hardware.
- El modelo de Machine Learning fue testeado por el investigador, realizando 20 repeticiones por cada gesto. Cada prueba se realizó alternando el número de k vecinos y el número de frames que se obtiene en la captura de datos con el sensor LMC. Obteniendo un porcentaje de exactitud del 100% en el gesto 1 (Open Hand), el 85% en el gesto 2 (Close Hand), el 65% en el gesto 3 (Wave In), el 95% en el gesto 4 (Wave Out), y el 10% en el gesto 5 (Pinch). En este sentido, se limita el funcionamiento del gesto 5, debido a su bajo porcentaje de exactitud.
- El mundo virtual, es decir, el modelo 3D que representa las instalaciones de la UTA fue generado y probado, controlando el avatar dentro del mundo con y sin elevaciones. Obteniendo que dentro de un modelo 3D generado sin elevaciones, el número máximo de FPS obtenido en áreas donde existen pocas edificaciones fue de 11FPS, y un máximo de 130FPS en áreas donde se concentra la mayor cantidad de edificaciones. El número de FPS influye en la velocidad con la que se cargan las imágenes y se muestran al usuario. Es decir, mientras mayor sea el número de FPS, más realista y fluida se verá la imagen. Además, se pudo evidenciar durante las pruebas de la plataforma que, a pesar de recibir instrucciones y demostraciones por parte del investigador, los usuarios presentan problemas para controlar el

avatar. Estos problemas radican en la dificultad que tienen para emplear los gestos de forma correcta.

4.2. Recomendaciones

- Para un mejor funcionamiento de la plataforma, se recomienda utilizar un computador cuyas características (memoria RAM, procesador y generación) sean superiores a las que se muestran en la **Tabla 7**. En vista que al ejecutar sobre el mismo computador tanto el modelo de reconocimiento de gestos, como la plataforma, producen retrasos en la reacción del avatar y ocasionan problemas por falta de memoria en el computador.
- Se recomienda analizar el desarrollo del modelo 3D, con el objetivo de mejorar el rendimiento del aplicativo en zonas con gran cantidad de componentes.
- Desarrollar un sistema de entrenamiento de gestos, donde el usuario pueda conocer los gestos y adaptarse al movimiento de la mano, antes de empezar con el recorrido virtual.
- Se recomienda a futuro, escalar la aplicación no solo para el ámbito educativo, sino para actividades como turismo dentro de la ciudad.

BIBLIOGRAFÍA

- [1] R. Nautiyal and R. Polus, “Virtual tours as a solidarity tourism product?,” *Annals of Tourism Research Empirical Insights*, vol. 3, no. 2, p. 100066, Nov. 2022, doi: 10.1016/j.annale.2022.100066.
- [2] Organización Mundial de la Salud, “COVID-19: cronología de la actuación de la OMS,” Apr. 27, 2020.
- [3] K. S. Dhanasree, K. K. Nisha, and R. Jayakrishnan, “Hospital Emergency Room Training Using Virtual Reality and Leap Motion Sensor,” in *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, IEEE, Jun. 2018, pp. 924–928. doi: 10.1109/ICCONS.2018.8662900.
- [4] F. Andrade, “Un enfoque inteligente para el reconocimiento de gestos manuales.”
- [5] J. Octavio and I. Denis, “Test of a Myo Armband,” *Revista de Ciencias Ambientales y Recursos Naturales*, vol. 3, no. 10, pp. 48–56, Nov. 2017.
- [6] J. T. López and C. Sagüés, “Reconocimiento e interpretación de gestos con dispositivo Leap,” 2013.
- [7] R. Nogales and M. Benalcazar, “Reconocimiento de Gestos de la Mano en Tiempo Real Usando Leap Motion Controller y Machine Learning,” *Conference Proceedings*, vol. 3, no. 1, pp. 823–835, 2019.
- [8] S. Sumpeno, I. G. A. Dharmayasa, S. M. S. Nugroho, and D. Purwitasari, “Immersive Hand Gesture for Virtual Museum using Leap Motion Sensor Based on K-Nearest Neighbor,” in *2019 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, IEEE, Nov. 2019, pp. 1–6. doi: 10.1109/CENIM48368.2019.8973273.
- [9] Y. Mizuchi and T. Inamura, “Cloud-based multimodal human-robot interaction simulator utilizing ROS and unity frameworks,” in *2017 IEEE/SICE International Symposium on System Integration (SII)*, IEEE, Dec. 2017, pp.

948–955. doi: 10.1109/SII.2017.8279345.

- [10] A. Clark and D. Moodley, “A System for a Hand Gesture-Manipulated Virtual Reality Environment,” in *Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists on - SAICSIT '16*, New York, New York, USA: ACM Press, 2016, pp. 1–10. doi: 10.1145/2987491.2987511.
- [11] M. H. Weik, “human-machine interface,” in *Computer Science and Communications Dictionary*, Boston, MA: Springer US, 2000, pp. 738–738. doi: 10.1007/1-4020-0613-6_8515.
- [12] A. Panesar, “What Is Artificial Intelligence?,” in *Machine Learning and AI for Healthcare*, Berkeley, CA: Apress, 2019, pp. 1–19. doi: 10.1007/978-1-4842-3799-1_1.
- [13] A. Panesar, “What Is Machine Learning?,” in *Machine Learning and AI for Healthcare*, Apress, 2019, pp. 75–118. doi: 10.1007/978-1-4842-3799-1_3.
- [14] *Encyclopedia of Applied and Computational Mathematics*. Springer Berlin Heidelberg, 2015. doi: 10.1007/978-3-540-70529-1.
- [15] J. Zea, “IMPLEMENTACIÓN DE UN SISTEMA DE CLASIFICACIÓN DE GESTOS DEL BRAZO HUMANO UTILIZANDO MYO ARMBAND PARA MANDO A DISTANCIA DE UN BRAZO ROBÓTICO DE 3GDL,” Quito, Dec. 2017.
- [16] M. Łukasik, T. Kuãmierczyk, Ł. Bolikowski, and H. S. Nguyen, “Hierarchical, multi-label classification of scholarly publications: Modifications of ML-KNN algorithm,” *Studies in Computational Intelligence*, vol. 467, pp. 343–363, 2013, doi: 10.1007/978-3-642-35647-6_22.
- [17] X. F. Zhong, S. Z. Guo, L. Gao, H. Shan, and J. H. Zheng, “An improved k-NN classification with Dynamic K,” in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Feb. 2017, pp. 211–216. doi: 10.1145/3055635.3056604.

- [18] R. Nogales, M. E. Benalcazar, B. Toalumbo, A. Palate, R. Martinez, and J. Vargas, “Construction of a Dataset for Static and Dynamic Hand Tracking Using a Non-invasive Environment,” 2021, pp. 185–197. doi: 10.1007/978-981-33-4565-2_12.
- [19] R. Muñoz and C. Rusu, “Mundos virtuales: ¿usabilidad real?,” Porto Alegre, BRA: Brazilian Computer Society, 2011, pp. 134–138.
- [20] A. S. Praetorius, L. Krautmacher, G. Tullius, and C. Curio, “User-Avatar Relationships in Various Contexts,” in *Mensch und Computer 2021*, New York, NY, USA: ACM, Sep. 2021, pp. 275–280. doi: 10.1145/3473856.3474007.

Anexos

Anexo 1. Primera parte de la encuesta

<p>Nombre y Apellido</p> <p>Tu respuesta _____</p>
<p>¿Qué edad tiene? *</p> <p><input type="radio"/> Menor de 18 años</p> <p><input type="radio"/> 18 años</p> <p><input type="radio"/> Mayor a 18 años</p>
<p>¿Pertenece a la Universidad Técnica de Ambato? *</p> <p><input type="radio"/> Si</p> <p><input type="radio"/> No</p>
<p>¿Conoce las instalaciones del campus Huachi de la Universidad Técnica de Ambato? *</p> <p><input type="radio"/> Si</p> <p><input type="radio"/> No</p>
<p>¿Le interesaría el desarrollo de una plataforma de interacción virtual del campus Huachi de la Universidad Técnica de Ambato?</p> <p><input type="radio"/> Me interesa mucho</p> <p><input type="radio"/> Me interesa</p> <p><input type="radio"/> Tal vez</p> <p><input type="radio"/> No es de mi interés</p>
<p>¿Estaría dispuesto a usar una plataforma de interacción virtual, para conocer las instalaciones del campus Huachi de la Universidad Técnica de Ambato?</p> <p><input type="radio"/> Si</p> <p><input type="radio"/> No</p>

Anexo 2. Segunda parte de la encuesta

Posterior al uso de la plataforma

Califique su experiencia utilizando la plataforma de interacción virtual del campus Huachi de la Universidad Técnica de Ambato? *

Muy buena

Buena

Regular

¿Volvería a usar la plataforma de interacción virtual, para conocer el campus Huachi de la Universidad Técnica de Ambato ? *

Si

No

¿Recomendaría el uso de la plataforma de interacción virtual, para conocer el campus Huachi de la Universidad Técnica de Ambato? *

Si

Tal vez

No

¿Tiene usted alguna sugerencia?

Tu respuesta _____