

e-ISSN 2550-6730

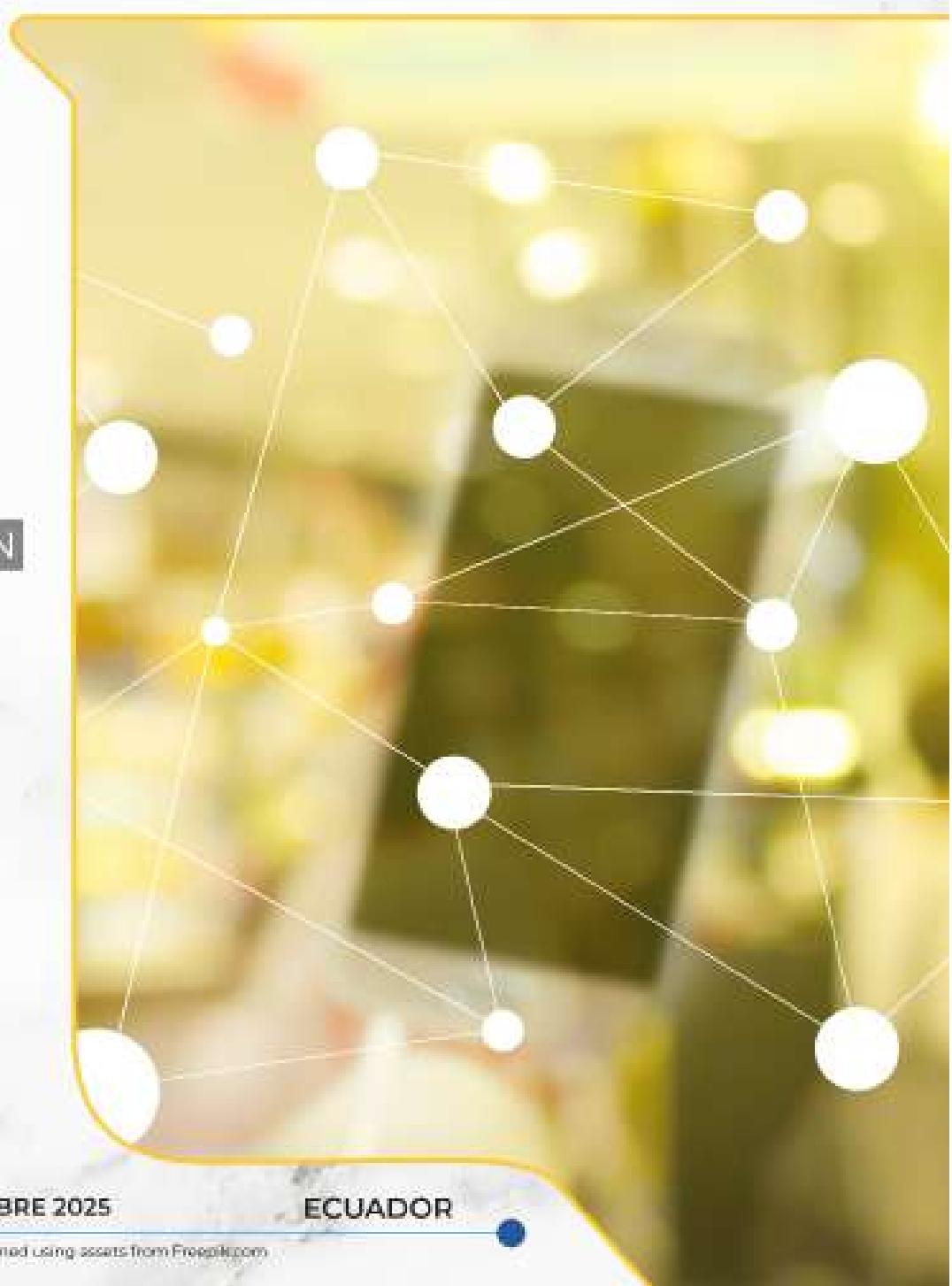


Revista de Tecnologías de la Informática  
y las Comunicaciones



UNIVERSIDAD  
TÉCNICA DE  
MANABÍ  
Fundada en 1952

**9**  
VOLUMEN  
Núm. 2



JUNIO - DICIEMBRE 2025

ECUADOR

This cover has been designed using assets from Freepik.com

## AUTORIDADES INSTITUCIONALES (INSTITUTIONAL AUTHORITIES)

*Rector*

**Santiago Quiroz Fernández, Ph. D.**

*Vicerrectora Académica*

**Mara Molina de Lozano, Ph. D.**

*Director de Investigación*

**Alex Dueñas Rivadeneira, Ph. D.**

*Decano de la Facultad de Ciencias Informáticas*

**Jorge Veloz Zambrano, Mg.**

## CONSEJO EDITORIAL (EDITORIAL BOARD)

*Director de la Revista*

*(Editor in Chief)*



**Jorge Párraga Álava, Ph.D.**



Universidad Técnica de Manabí, Ecuador

## EDITORES (Editors)

**Leticia Vaca Cárdenas, Ph. D.**

Universidad Técnica de Manabí, Ecuador

**Lucia Rivadeneira Barreiro, Ph. D.**

Universidad Técnica de Manabí, Ecuador

**Leonardo Chancay García, Ph. D**

Universidad Técnica de Manabí, Ecuador

**Maricela Pinargote Ortega, Ph.D**

Universidad Técnica de Manabí, Ecuador

## CONSEJO DE REVISORES (REVIEWERS BOARD)

 **Ph.D. Diego Hernán Peluffo Ordóñez**  
 Université Mohammed VI Polytechnique, Marruecos

 **Ms.C. Alcibar Yáñez Escobar**  
 Universidad de Málaga, España

 **Ph.D. Alex Santamaría Philco**  
 Universitat Politècnica de València, España

 **Ph.D. Marlon Navia Mendoza**  
 Universidad Técnica de Manabí, Ecuador

 **Ms. C. Wilmer Moreira Sánchez**  
 Universidad Técnica de Manabí, Ecuador

 **Ms. C. Tatiana Zambrano Zambrano**  
 Universidad Técnica de Manabí, Ecuador

 **Ms. C. Enrique Macias Arias**  
 Universidad Técnica de Manabí, Ecuador

 **Ms. C. Gabriel Morejón López**  
 Universidad Técnica de Manabí, Ecuador

## CONSEJO CIENTÍFICO (ADVISORY BOARD)

 **Felipe Bello Robles, Ph. D.**  
 Universidad de Santiago de Chile, Chile

 **Manuel Villalobos Cid, Ph. D.**  
 Universidad de Santiago de Chile, Chile

 **Paulo Freitas de Oliveira Novais, Ph. D.**  
 Univerdidade do Minho, Portugal

 **Dalila Alves Durães, Ph. D.**  
 Univerdidade do Minho, Portugal

 **Edith Josefina Liccioni, Ph. D.**  
 Universidad de Chimborazo, Ecuador

 **Cristóbal Samaniego Alvarado, Ph. D.**  
 Barcelona Supercomputing Center, España

## EQUIPO TÉCNICO (TECHNICAL TEAM)

*Webmaster OJS*  
 **Ing. Victor López Tuárez**  
Dirección de Investigación,  
Universidad Técnica de Manabí, Ecuador

*Diseñador, Diagramación y Portada*  
 **Ing. Irving Cevallos Bumbila, Mg.**  
Dirección de Comunicaciones,  
Universidad Técnica de Manabí, Ecuador

**Informática y Sistemas:**  
**Revista de Tecnologías de la Informática y las Comunicaciones**  
**Volumen 9, Número 2**  
**Julio – Diciembre 2025**  
**e-ISSN: 2550-6730**

Informática y Sistemas: Revista de Tecnologías de la Informática y las Comunicaciones (ISRTIC) es una publicación electrónica semestral de carácter científico, que edita la Facultad de Ciencias Informáticas de la Universidad Técnica de Manabí, orientada a la socialización de resultados de investigación, a través de artículos novedosos y de alto rigor científico, en las áreas asociadas a las tecnologías de la información y las comunicaciones. ISRTIC no efectúa cargos por concepto de costos de procesamiento, envío o publicación de artículos.

*El proceso editorial de ISRTIC se gestiona a través del*



*ISRTIC es una publicación de acceso abierto con licencia*



*Los artículos de ISRTIC cuentan con código de identificación de objeto digital (DOI)*



*Los autores de ISRTIC usan un Identificador Abierto de Investigador y Colaborador (ORCID)*



*ISRTIC utiliza el sistema antiplagio*



*Las revista está indizada en*



Google académico

DOAJ DIRECTORY OF OPEN ACCESS JOURNALS



*Los artículos de la presente edición se pueden obtener en  
<https://doi.org/10.33936/isrtic.v9i2>*

## **EDITORIAL**

Lectoras/es,

El presente número de Informática y Sistemas marca un hito significativo para nuestra revista: *la incorporación oficial a Redalyc*, una de las plataformas de mayor prestigio en el ecosistema de la ciencia abierta iberoamericana. Este logro consolida el compromiso sostenido con la calidad editorial, la rigurosidad científica y la visibilidad internacional de la producción académica que aquí se publica. Agradecemos a la comunidad de investigadores, evaluadores y equipos institucionales que han contribuido a alcanzar este importante espacio de indexación.

Los artículos que conforman este volumen reflejan la diversidad y actualidad de las líneas de investigación desarrolladas en instituciones como la Universidad Técnica de Machala, la Universidad de Vigo, la Escuela Superior Politécnica Agropecuaria de Manabí “Manuel Félix López”, la Universidad Nacional de Chimborazo y la Universidad Estatal Península de Santa Elena. Sus aportes muestran el avance regional en temas emergentes de ingeniería, ciencia de datos, tecnologías aplicadas y transformación digital.

El número inicia con una revisión sobre el uso de la realidad aumentada en dispositivos móviles para la educación, seguida de un estudio sobre generación y validación de guiones educativos con inteligencia artificial. Continúa con un análisis de diseño centrado en el usuario para aplicaciones web y un aporte en ciberseguridad sobre buenas prácticas en la implementación de soluciones de IA en la nube. El volumen incluye además una propuesta de balanceo de datos con redes generativas en el ámbito agrícola, junto con iniciativas de monitoreo ganadero mediante IoT, herramientas conversacionales para orientación vocacional y sistemas inteligentes para la detección en tiempo real de actividades u objetos sospechosos.

Estas contribuciones reflejan la vitalidad investigativa de las universidades y su capacidad para producir conocimiento pertinente y de alcance regional, reforzando el compromiso de esta edición con la difusión científica de alto impacto y el fortalecimiento de un ecosistema académico más articulado y visible.

Invitamos a la comunidad a continuar fortaleciendo este espacio editorial, contribuyendo con investigaciones que impulsen el avance científico y tecnológico en beneficio de la sociedad.

Atentamente,

**Jorge Párraga Álava, Ph.D.**

Editor en Jefe

Informática y Sistemas:

Revista de Tecnologías de la Información y Computación

# ÍNDICE

104-125

**Revisión sobre el Uso de la Realidad Aumentada en Dispositivos Móviles en el Ámbito Educativo**

*Review on the Use of Augmented Reality in Mobile Devices in the Educational Field*

Jeniffer Karina Alcívar Pazmiño, Yimmy Salvador Loor Vera , Joffre Edgardo Panchana Flores

126-139

**La Innovación Educativa con Inteligencia Artificial: Generación y Validación de Guiones para Monólogos Educativos en la Enseñanza Universitaria**

*Education Innovation with Artificial Intelligence: Generation and Validation of Scripts for Educational Monologues in University Teaching*

Christiam Xavier Núñez Zavala, Jorge Noé Silva Castillo, Cristhy Nataly Jimenez Granizo, Hernán Ramiro Pailiacho Yucta

140-151

**Diseño Basado en el Usuario en Aplicaciones Web**

*User-Centered Design in Web Applications*

Miriam Lizeth Lectong Anchundia, Jéssica Johanna Morales Carrillo

152-163

**Buenas Prácticas de Seguridad para la Implementación de Inteligencia Artificial en Entornos de Computación en la Nube**

*Practices for Securing the Implementation of Artificial Intelligence in Cloud-Based Environments*

Yasbeck Jemima Mora Chavez, Reinaldo Benedicto Fernández Gonzalez, Joofre Antonio Honores Tapia, Milton Rafael Valarezo Pardo

164-176

**Balanceo de Conjuntos de Datos Basado en Redes Generativas Aplicado a Imágenes del Sector Agrícola**

*Generative Network-Based Data Balancing Applied to Agricultural Image Datasets*

Luis Jesús Montesdeoca Espinoza, Stalin Joel Zambrano Rojas, Victor Joel Pinargote Bravo, Luis Cristobal Cedeño Valarezo

**177-183**

**Monitoreo en tiempo real del ganado: Integración de LoRa, sensores y visualización de datos**

*Real-time monitoring of cattle: integration of LoRa, sensors, and data visualization*

Walter Orozco Iguasnía, Juleidy Jailin Tigrero Muñoz, Santiago Andrés Orozco Villarroe

**184-196**

**Desarrollo de un Chatbot con Inteligencia Artificial para orientación vocacional según el perfil competencial de estudiantes de tercero bachillerato**

*Development of a Chatbot with Artificial Intelligence for Career Guidance Based on the Competency Profile of Third-Year High School Students*

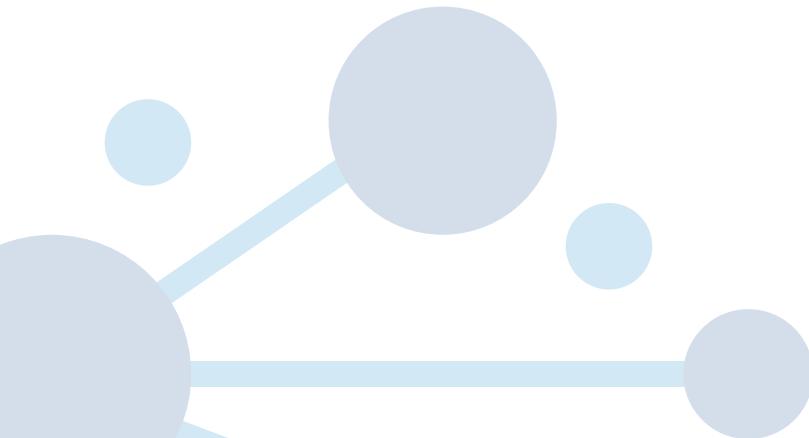
Fernando Ericsson Montoya Naguas, Carlos Elian Lalangui Flores, Fausto Fabián Redrován Castillo, Freddy Aníbal Jumbo Castillo

**197-213**

**Prototipo de Sistema de Monitoreo Inteligente para la detección en Tiempo Real de Objetos y Actividades Sospechosas Utilizando Deep Learning**

*Prototype of an Intelligent Monitoring System for Real-Time Detection of Suspicious Objects and Activities Using Deep Learning*

Lauro Alfonso Erreyes Cuenca, Nahin Josue Olmedo Chica, Mariuxi Paola Zea Ordóñez, Nancy Magaly Loja Mora





## Revisión sobre el Uso de la Realidad Aumentada en Dispositivos Móviles en el Ámbito Educativo

### Review on the Use of Augmented Reality in Mobile Devices in the Educational Field

#### Autores

\* **Jeniffer Karina Alcivar Pazmiño**

✉ jenifferk.alcivar@espm.edu.ec

**Loor Vera Yimmy Salvador**

✉ yloor@espm.edu.ec

**Panchana Flores Joffre Edgardo**

✉ edgardo.panchana@uleam.edu.ec

#### Resumen

El uso de la realidad aumentada (RA) en dispositivos móviles constituye un recurso de alto impacto en la transformación del aprendizaje en experiencias más inmersivas. Esta investigación tiene por objetivo indagar sobre el uso de la RA en el ámbito educativo, tomando como base artículos científicos publicados entre 2020 y 2024, clasificando casos de uso, áreas de aplicación, limitaciones y beneficios reportados en estudios recientes. Se siguió la metodología de revisión sistemática propuesta por Kitchenham en sus tres fases: planificación, realización y documentación de la revisión, combinada con el método analítico-sintético que facilitó el análisis minucioso de las fuentes obtenidas a través de la revisión bibliográfica. En la fase de planificación se definieron las preguntas de investigación, los criterios de búsqueda, selección de palabras clave y bases de datos científicas; en la fase de realización, se obtuvieron 793 artículos, de los cuales se seleccionaron 50 que cumplieron con los criterios establecidos; para la fase de documentación se diseñó una tabla con campos específicos donde se recogió y organizó la información relevante extraída de cada artículo. Los resultados evidencian que el uso de la RA en dispositivos móviles facilita el aprendizaje mediante la visualización de modelos tridimensionales y contenido interactivo, fomentando la comprensión de conceptos complejos (20%) en áreas curriculares como matemáticas (26%) y química (20%), motivando y fortaleciendo el compromiso de los estudiantes (70%); no obstante, la falta de formación en los docentes (60%) y la necesidad de una adecuada infraestructura tecnológica (54%) limitan su implementación.

**Palabras clave:** Realidad aumentada; Aprendizaje interactivo; Tecnología educativa; Aplicaciones Móviles; Entornos de Aprendizaje.

<sup>1</sup> Escuela Superior Politécnica Agropecuaria de Manabí “Manuel Félix López”, Calceta, Manabí, Ecuador.

<sup>2</sup> Universidad Laica Eloy Alfaro de Manabí, Manta, Manabí, Ecuador .

\*Autor para correspondencia

#### Abstract

The use of augmented reality (AR) on mobile devices represents a high-impact resource in transforming learning into more immersive experiences. This study aims to explore the use of AR in educational contexts by reviewing scientific articles published between 2020 and 2024, categorizing use cases, application areas, reported limitations, and benefits. A systematic literature review was conducted following Kitchenham's three-phase methodology: planning, conducting, and reporting, complemented by an analytical-synthetic method to support a detailed examination of the selected sources. During the planning phase, research questions were defined along with search criteria, relevant keywords, and scientific databases. In the conducting phase, a total of 793 articles were retrieved, from which 50 were selected based on predefined inclusion criteria. For the reporting phase, a data extraction table was designed to collect and organize relevant information from each study. The results indicate that mobile AR enhances learning through the visualization of three-dimensional models and interactive content, supporting the understanding of complex concepts (20%) in subject areas such as mathematics (26%) and chemistry (20%), while also promoting student motivation and engagement (70%). However, limitations persist, particularly regarding insufficient teacher training (60%) and the need for appropriate technological infrastructure (54%).

**Keywords:** Augmented Reality; Interactive Learning; Educational Technology; Mobile Applications; Learning Environments.

Enviado: 21/04/2025

Aceptado: 25/06/2025

Publicado: 02/07/2025



## 1. Introducción

La tecnología ha transformado significativamente el contexto educativo, gracias a que ofrece múltiples herramientas que potencian el aprendizaje interactivo y personalizado.

Una de la más destacadas es la realidad aumentada (RA), tecnología emergente que permite integrar elementos virtuales en el entorno real para crear experiencias educativas inmersivas (Kamińska et al., 2023).

Esta integración hace posible que los estudiantes interactúen con contenido tridimensional, visualizaciones dinámicas y simulaciones prácticas, fomentando un aprendizaje multisensorial (Aguilar Acevedo et al., 2023); gracias a su versatilidad, se puede aprovechar en múltiples áreas educativas. Según González Pérez & Cerezo Cortijo (2020) la implementación de RA en el área de ciencias ha contribuido en la comprensión conceptos abstractos, mediante simulaciones interactivas, por parte de los estudiantes, disminuyendo así la carga cognitiva en el desarrollo de actividades que pueden considerar complejas.

El uso de la RA en el contexto educativo global se ha incrementado exponencialmente gracias a la posibilidad de implementarla en todo tipo de dispositivos móviles, que al ser altamente accesibles y estar equipados con cámaras y sensores, se han constituido como la plataforma ideal para implementar soluciones que, además de ser económicas, son altamente portátiles y versátiles (Leal Aragón, 2020), potenciando su uso en el aprendizaje móvil de alguna disciplinas como por ejemplo las matemáticas (O. M. Martínez et al., 2021).

En este contexto, Márquez Díaz & Morales Espinosa (2020) aseguran que la utilización de aplicaciones basadas en códigos QR y modelos 3D, han facilitado la enseñanza de funciones algebraicas y trascendentales y, de esta manera, mejorado la interacción y comprensión de estos contenidos por parte de los estudiantes dentro y fuera del aula.

Por su parte, para González Pérez & Cerezo Cortijo (2020) la RA fomenta el aprendizaje activo al hacer posible la transformación de las dinámicas tradicionales en experiencias más colaborativas e innovadoras que promueven entornos que potencian la participación y la interacción significativa.

Pese a sus numerosas ventajas, el uso de la RA en dispositivos móviles en el ámbito educativo enfrenta desafíos significativos relacionados con su adopción, costos de los smartphones, implementación efectiva y evaluación de su impacto pedagógico (Cabascango Trávez, 2023).

Núñez Zavala et al. (2023) mencionan, en su estudio desarrollado en unidades educativas del sector rural, que carecen de medios

económicos para adquirir recursos tecnológicos que permitan apoyar su proceso de enseñanza-aprendizaje en sus estudiantes. A pesar de estas y otras dificultades más que se pueden encontrar, los beneficios que aporta superan las barreras tecnológicas y formativas.

Por ello, resulta fundamental realizar un análisis sistemático que identifique las tendencias actuales, las áreas de aplicación más estudiadas, las metodologías empleadas y los resultados obtenidos en investigaciones previas (Calderón Imbaquingo et al., 2024).

Sin embargo, a pesar del creciente interés y la exploración del potencial de la RA móvil en la educación, existe una brecha explícita en el conocimiento sobre las áreas donde la RA tiene mayor impacto educativo, las limitaciones que enfrenta y los beneficios que tiene el uso de esta tecnología para mejorar el aprendizaje.

La revisión sistemática que brinda este estudio a diferencia de investigaciones previas es su análisis transversal más amplio y actualizado, abarcando diferentes niveles educativos y contextos pedagógicos que permiten comprender el desarrollo de la realidad aumentada en determinados entornos y situaciones.

El uso de esta tecnología está en constante evolución, lo cual es fundamental esta revisión con datos recientes que identifica nuevas oportunidades de aplicación, detectando sus barreras y mejoras para el aprendizaje.

El objetivo de este artículo es identificar los casos de uso más frecuentes, áreas de aplicación, beneficios reportados y principales limitaciones encontradas sobre el uso de la realidad aumentada en dispositivos móviles en contextos educativos entre 2020 y 2024. Para ello, se aplicó la metodología de revisión sistemática propuesta por Kitchenham estructurada en tres fases clave. En la fase de planificación, se formulan cuatro preguntas de investigación que permitieron delimitar el enfoque del estudio; en la fase de ejecución, se realiza la búsqueda y selección de estudios relevantes, además de la extracción de datos esenciales; por último, en la fase de documentación, se organizaron y, posteriormente, se analizaron los resultados obtenidos, mismos que ofrecen una visión actualizada sobre el uso de la realidad aumentada en los dispositivos móviles en el ámbito educativo. Las preguntas de investigación y los objetivos relacionados que guían el desarrollo del estudio se muestran en la tabla 1.

## 2. Materiales y Métodos

Se aplicó el método analítico-sintético para descomponer y examinar en detalle las fuentes recopiladas. Este enfoque metodológico facilitó la identificación de patrones y aportes

**Tabla 1.** Preguntas de investigación y objetivos relacionados. Fuente: Los autores.

| Pregunta de investigación  | Objetivo relacionado   |
|--|--|
| ¿Cuáles son los principales casos de uso de la realidad aumentada en dispositivos móviles dentro del ámbito educativo? | Analizar los principales casos de uso de la realidad aumentada en dispositivos móviles dentro del ámbito educativo |
| ¿En qué áreas del currículo educativo se ha encontrado un mayor impacto de la RA?                                      | Identificar las áreas del currículo educativo en las que esta tecnología ha tenido mayor impacto                   |
| ¿Qué limitaciones enfrentan las aplicaciones de RA en dispositivos móviles en el ámbito educativo?                     | Examinar las limitaciones más comunes que enfrentan estas aplicaciones en contextos educativos                     |
| ¿Qué beneficios se han identificado en el uso de esta tecnología para mejorar el aprendizaje en los últimos 5 años?    | Evaluuar los beneficios pedagógicos que se han reportado durante los últimos cinco años                            |

clave en las fuentes bibliográficas revisadas, los cuales fueron posteriormente sintetizados con el fin de destacar los hallazgos más relevantes y dar respuesta al objetivo de este estudio.

Otro método utilizado fue el de revisión sistemática de Kitchenham, ampliamente reconocido y utilizado en el campo de la ingeniería de software y muchas otras disciplinas, gracias a su capacidad para garantizar una evaluación exhaustiva y confiable de la literatura disponible (Guaña Moya, 2024); este método proporcionó un enfoque riguroso y estructurado para la identificación, evaluación y síntesis de la evidencia existente sobre el uso de la realidad aumentada en dispositivos móviles en el ámbito educativo (Silva & Santana, 2024).

Según Tebes et al. (2019), el método de Kitchenham fue propuesto originalmente en 2004 y ha sido actualizado en diversas publicaciones posteriores, especialmente en 2007, cuando se publicó el “Guidelines for performing Systematic Literature Reviews in Software Engineering”; resaltando que, para la ejecución de una revisión sistemática de literatura de debe desarrollar a través de tres fases que son:

- Planificación de la revisión
- Realizar la revisión
- Documentar la revisión

Al respecto, Ali et al. (2022) menciona que las etapas correspondientes a cada fase de una revisión sistemática pueden variar dependiendo del contexto de la investigación, permitiendo así a los investigadores centrarse en aquellas etapas que aporten mayor valor a los objetivos específicos de su estudio, sin comprometer la validez de los resultados.

De manera similar, Carrión y Serrano (2021) resaltan que la flexibilidad en la aplicación de estas fases permite adaptar el proceso a las necesidades específicas del estudio, garantizando así un equilibrio entre rigurosidad y practicidad. A continuación, se presentan las fases y etapas que fueron consideradas en el estudio:

#### Fase 1: Planificación de la revisión

Se centra en establecer las bases para una revisión sistemática de

la literatura. En esta fase se define el propósito y la necesidad de la revisión.

También se formulan las preguntas de investigación que guiarán el desarrollo del estudio y se los criterios de inclusión y exclusión, así como las estrategias de búsqueda (Sánchez Hayman et al., 2024).

#### Identificación de la necesidad de la revisión

La necesidad de este estudio se basa en el creciente interés respecto de la implementación de realidad aumentada (RA) en dispositivos móviles en el ámbito educativo, para comprender el estado actual de la investigación en este campo.

#### Preguntas de investigación

Para delimitar la investigación se formularon las siguientes preguntas:

- **PI1:** ¿Cuáles son los principales casos de uso de la realidad aumentada en dispositivos móviles dentro del ámbito educativo?
- **PI2:** ¿En qué áreas del currículo educativo se ha encontrado un mayor impacto de la RA?
- **PI3:** ¿Qué limitaciones enfrentan las aplicaciones de RA en dispositivos móviles en el ámbito educativo?
- **PI4:** ¿Qué beneficios se han identificado en el uso de esta tecnología para mejorar el aprendizaje en los últimos 5 años?

#### Desarrollo del protocolo de revisión

Se elaboró un protocolo para garantizar la transparencia del estudio, el cual incluyó:

Criterios de inclusión:

- **Año de publicación:** Entre 2020 y 2024, para asegurar una revisión con base en los estudios más recientes.
- **Contexto:** Los estudios deben estar relacionados directamente con la realidad aumentada en dispositivos móviles con un enfoque educativo.
- **Contenido:** La documentación encontrada debe abordar



aspectos clave como aplicaciones, beneficios y nuevas tecnologías de la RA en el ámbito educativo.

#### Criterios de exclusión:

- **Estudios desactualizados:** Se excluyeron estudios publicados antes de 2020, ya que no representarían un valor significativo para este estudio cuyo eje es el ámbito tecnológico.
- **Otros contextos:** Estudios que abordan la realidad aumentada en dispositivos móviles pero que no están relacionados con el contexto educativo, fueron excluidos.
- **Fuente no válidas:** Se descartaron estudios considerados como no académicos: artículos de blogs, fuentes no científicas e informes técnicos sin revisiones o considerados no confiables.

#### Base de datos:

Se seleccionaron las bases de datos Google Scholar, IEEE Xplore, Redalyc, ScienceDirect, Scielo, Scopus y WoS (Web of Science), que garantizan el acceso a estudios actualizados, de gran valor académico.

Estas plataformas permiten realizar una revisión exhaustiva de la literatura al incluir investigaciones de alta calidad y enfoques variados, esenciales para abordar los objetivos del estudio (Sacoto et al., 2024).

#### Estrategia de la búsqueda:

Se diseñó una estrategia de búsqueda basada en la definición y combinación de palabras clave para cada una de las bases de datos seleccionadas, como se muestra en la tabla 2, para poder identificar artículos relevantes que abarquen conceptos fundamentales relacionados con la realidad aumentada en el ámbito educativo y su implementación en dispositivos móviles.

**Tabla 2.** Palabras clave utilizadas para la revisión de literatura.

Fuente: Los autores.

| Bases de datos | Palabras claves utilizadas   |
|----------------|--|
| Google Scholar | “Realidad aumentada en educación”, “Realidad aumentada en dispositivos móviles”, “Uso pedagógico de la realidad aumentada”, “Tecnologías educativas y realidad aumentada”. |
| IEEE Xplore    | “Augmented Reality in Education”, “Augmented reality in mobile learning”, “Interactive learning tools with AR” “Pedagogical use of augmented reality”                      |
| Redalyc        | “Realidad aumentada en educación”, “Realidad aumentada en dispositivos móviles”, “Aplicaciones pedagógicas de la realidad aumentada”                                       |
| ScienceDirect  | “Augmented Reality in Education”, “Augmented reality in mobile learning”, “Interactive learning tools with AR” “Pedagogical use of augmented reality”                      |
| Scielo         | “Realidad aumentada en dispositivos móviles”, “Realidad aumentada y aprendizaje móvil”, “Innovaciones tecnológicas en educación con RA”                                    |
| Scopus         | “Augmented Reality in Education”, “Augmented reality in mobile learning”, “Interactive learning tools with AR” “Pedagogical use of augmented reality”.                     |
| WoS            | “Augmented Reality in Education”, “Augmented reality in mobile learning”, “Interactive learning tools with AR” “Pedagogical use of augmented reality”                      |



**Figura 1.** Proceso para sintetizar los datos.

Fuente: Los autores.

Se obtuvieron 793 artículos relacionados, los mismos que luego de pasar por el proceso de evaluación, resultaron en 50 artículos, los cuales contenían la información necesaria para el desarrollo de la revisión sistemática de literatura.

### Extracción de datos

Se extrajo información clave de cada estudio seleccionado, con el propósito de organizar y sintetizar de manera efectiva los datos relevantes para la revisión sistemática. La tabla 3 muestra el detalle de los campos definidos para la organización de los datos extraídos, asegurando una evaluación integral y consistente en el posterior análisis y comparación de los estudios seleccionados.

### Sintetizar los datos

Con los 50 artículos seleccionados para esta fase, se procedió a realizar un análisis sistemático que permitió identificar patrones, tendencias y obtener las respectivas respuestas a cada una de las preguntas de investigación. Este proceso se realizó de la siguiente manera:

### Fase 3: Documentar la revisión

Consiste en la elaboración del informe final de la revisión sistemática, en el que se presentan de forma clara y estructurada los resultados obtenidos de la revisión sistemática. Tal como lo

**Tabla 3.** Campos considerados en la recolección de información.

Fuente: Los autores.

| Campos       | Descripción   |
|--------------|---|
| Título       | Título del artículo.  |
| Autor / Año  | Nombre del autor o autores del artículo y el año de publicación.                    |
| Casos de uso | ¿Cómo se utiliza la realidad aumentada en el contexto educativo?                    |
| Áreas        | Áreas del currículo educativo en las que se aplica la RA.                           |
| Limitaciones | Restricciones o desafíos encontrados en la implementación de la realidad aumentada. |
| Beneficios   | Contribución o beneficios del uso de la realidad aumentada al proceso educativo.    |

manifiesta Arellano Pimentel et al. (2024) en esta fase se organiza la información para responder las preguntas de investigación planteadas, destacando los hallazgos más relevantes.

## 3. Resultados y Discusión

### Identificación y selección de estudios

En la primera selección, se aplicó el criterio año de publicación, asegurando que la revisión se base en los avances tecnológicos recientes entre 2020 y 2024; como resultado, de los 793 documentos inicialmente identificados, se preseleccionaron 589 que cumplieron con el filtro aplicado. En la segunda selección, se aplicó el criterio de inclusión referente al contexto; en este sentido, se excluyeron aquellos documentos que, aunque trataban sobre realidad aumentada en dispositivos móviles, no estaban relacionados con el ámbito educativo, disminuyendo el total de artículos seleccionados a 374. Finalmente, en la tercera selección, se aplicó el criterio de inclusión con base al contenido completo del documento, incluyendo aquellos estudios que analizaban aplicaciones, beneficios y nuevas tecnologías de la realidad aumentada en el contexto educativo y excluyendo fuentes no académicas, como artículos de blogs, informes técnicos sin revisión y documentos de fuentes no científicas, para garantizar la calidad y validez de la evidencia recopilada. De esta manera, se seleccionaron 50 artículos relevantes y pertinentes para la revisión sistemática.

En la tabla 4 se presentan los resultados de la búsqueda realizada en las diferentes bases de datos científicas, utilizando las palabras claves definidas en la tabla 2, así como la aplicación de los criterios de inclusión y exclusión establecidos.

Se muestran los artículos seleccionados en tres etapas de selección, de acuerdo con el proceso y metodología diseñada.

A partir de la cantidad de documentos encontrados, se analizó la evolución de las publicaciones sobre el uso de la realidad aumentada en dispositivos móviles dentro del ámbito educativo. La figura 2 muestra esta tendencia, evidenciando un incremento significativo en los últimos cinco años. Este crecimiento refleja el creciente interés dentro de la comunidad científica, así como también su relevancia en el campo de la educación.

### Extracción de los datos

La información sintetizada se presenta en la tabla 5, que incluye los campos definidos en la tabla 3. Este enfoque permitió responder de manera efectiva cada una de las preguntas de investigación (PI) definidas durante la fase de planificación del proceso de revisión sistemática.



**Tabla 4.** Documentos encontrados y seleccionados.  
Fuente: Los autores.

| Bases de datos | Cant. encontrados | Criterios de exclusión/inclusión |                |                 |
|----------------|-------------------|----------------------------------|----------------|-----------------|
|                |                   | 1era. selección                  | 2da. selección | 3era. selección |
| Google Scholar | 102               | 94                               | 57             | 7               |
| IEEE Xplore    | 85                | 85                               | 45             | 6               |
| Redalyc        | 62                | 60                               | 32             | 6               |
| ScienceDirect  | 93                | 87                               | 69             | 6               |
| Scielo         | 49                | 47                               | 37             | 5               |
| Scopus         | 168               | 121                              | 71             | 11              |
| WoS            | 134               | 95                               | 63             | 11              |
| Total          | 793               | 589                              | 374            | 50              |

**Figura 2.** Documentos encontrados por año.  
Fuente: Los autores.



### Sintetizar los datos

Se examinaron de manera individual los campos relacionados con cada una de las preguntas de investigación (PI), asegurando que cada una de ellas sea abordada de forma exhaustiva, contribuyendo a la robustez y claridad de los resultados obtenidos.

### PI1: ¿Cuáles son los principales casos de uso de la realidad aumentada en dispositivos móviles dentro del ámbito educativo?

Se identificaron varios casos de uso de la realidad aumentada en el ámbito educativo, los que se muestran de manera sintetizada en la tabla 6, incluyendo su frecuencia de aparición, porcentaje (respecto al total de estudios analizados) y los artículos en los que se mencionan. De estos resultados, se identificó un amplio abanico de aplicaciones de la realidad aumentada en el ámbito educativo. La visualización de conceptos en 3D es la aplicación más frecuente con un equivalente del 20%, facilitando la representación de objetos y estructuras complejas en el aprendizaje. Le sigue la exploración de entornos virtuales,

con un 12%, permitiendo interactuar con escenarios simulados que enriquecen la comprensión contextual de determinados contenidos.

Por otra parte, las simulaciones interactivas representan un 10%, lo que sugiere que facilita representación dinámica de procesos y situaciones que enriquecen el aprendizaje práctico y experimental. Le sigue los juegos educativos, con un 8 %, los cuales fomentan la motivación y el aprendizaje mediante experiencias participativas. Así como también, aplicaciones en educación STEM, aprendizaje basado en proyectos y laboratorios virtuales, representando también un 8% de frecuencia, lo que evidencia su potencial para diversificar metodologías de enseñanza y enriquecer la experiencia educativa en distintos niveles y disciplinas.

Finalmente, aunque con menor frecuencia, con un porcentaje del 4%, se encontraron usos de la RA en la enseñanza de química, historia y patrimonio cultural, formación técnica y profesional, y educación inclusiva y personalizada.

Estos datos y cifras demuestran que la RA es una herramienta valiosa no solo en ciencias y matemáticas, sino que también en diversas áreas del conocimiento, ampliando sus beneficios en la educación.

Los casos de uso encontrados evidencian la amplia gama de aplicaciones de la RA en dispositivos móviles dentro del ámbito educativo. Desde la mejora de la comprensión de conceptos complejos hasta la creación de experiencias inmersivas y colaborativas, la RA se presenta como una herramienta poderosa para transformar los métodos de enseñanza y aprendizaje en diversos contextos educativos.

Esto responde a la pregunta de investigación sobre los principales casos de uso de la realidad aumentada, al demostrar cómo esta tecnología se está adaptando a distintas necesidades y disciplinas dentro del sector educativo.

### PI2: ¿En qué áreas del currículo educativo se ha encontrado un mayor impacto de la RA?

A través del análisis de los artículos seleccionados para esta revisión, se han identificado múltiples disciplinas en las que la RA ha sido utilizada para mejorar la comprensión de conceptos, promover el aprendizaje interactivo y aumentar la motivación de los estudiantes. En la tabla 7 se presentan las áreas de currículo encontradas en los estudios analizados.

El análisis de los datos revela que la RA ha tenido una mayor presencia en disciplinas como Matemáticas con 26% y Química con 20%, lo que indica su gran potencial en disciplinas científicas. Mientras que la Tecnología educativa y Ciencias de

**Tabla 5.** Datos obtenidos de los artículos seleccionados

Fuente: Los autores.

| Nº | Título  | Autor/Año                      | Casos de uso  | Áreas   | Limitaciones  | Beneficios  |
|----|---|--------------------------------|---|---|---|---|
| 1  | Análisis sistemático sobre el uso de la Realidad Aumentada en Educación Infantil  | (Rivas Rebaque et al., 2021)   | Aprendizaje de idiomas de forma interactiva.<br>Asimilación de números a través de libros de cuentos interactivos.<br>Enseñanza de formas geométricas y mejora de habilidades espaciales.                 | Idiomas.<br>Matemáticas.<br>Ciencias.<br>Educación artística.                               | Problemas técnicos, recursos limitados en algunas instituciones educativas y falta de familiaridad de los docentes con la tecnología.   | La RA mejora el desarrollo cognitivo, social, lingüístico y motivacional de los niños   |
| 2  | Desarrollo del Aprendizaje Basado en Proyectos con Realidad Aumentada en educación secundaria para mejorar rendimientos en el aula de música. | (Rodríguez López, 2022)        | Desarrollo de un Aprendizaje Basado en Proyectos (ABP) utilizando Realidad Aumentada (RA).<br>Uso de una aplicación de Realidad Aumentada denominada Halo   | Geografía.<br>Música.<br>Historia.  | Limitaciones en el acceso a dispositivos móviles para algunos estudiantes.<br>Necesidad de formación docente.   | Resultados académicos positivos, con alto porcentaje de respuestas correctas en las evaluaciones de clases.<br>Aumento en la motivación y participación de los estudiantes. |
| 3  | Aprendizaje integrado y colaborativo de ciencias a través de la realidad aumentada en educación infantil                                      | (Moreno Fuentes et al., 2021)  | Uso de un zoo virtual utilizando realidad aumentada para la enseñanza de ciencias.<br>Uso de aplicaciones como Quiver, Aurasma, y Animals AR para visualización de conceptos en temas de los seres vivos. | Ciencias  | Problemas técnicos y falta de recursos tecnológicos para una adecuada implementación de herramientas de realidad aumentada.<br>Dificultades para adaptar la adaptación curricular de manera efectiva con la integración de la realidad aumentada. | Aumento en la motivación de aprender por parte de los estudiantes.<br>Mejor comprensión en temas complejos relacionados con los seres vivos.                                |
| 4  | El uso de Realidad Aumentada en la Enseñanza de Ciencias Naturales en Educación Básica.   | (Urbina López et al., 2024)    | Uso de simulaciones interactivas en los experimentos científicos.<br>Visualización de conceptos abstractos en 3D.   | Ciencias Naturales<br>Química<br>Biología<br>Ciencias Ambientales                           |   | Aumento de la motivación y el compromiso del estudiante.<br>Mejora en la comprensión de conceptos complejos.  |
| 5  | Perspectiva tecnopedagógica de la realidad aumentada en la educación  | (Aguilar Acevedo et al., 2023) | Aplicaciones de realidad aumentada móvil para mejorar la experiencia de aprendizaje.<br>Simulaciones aumentadas en la enseñanza de ciencias.  | Ciencias (Biología, Física, Química)<br>Matemáticas.<br>Ingeniería.<br>Artes y Humanidades. | Falta de experiencia en el uso de la tecnología por parte de alumnos y profesores.<br>Costos asociados a la tecnología y posibles problemas técnicos  | Mayor participación y motivación de los estudiantes en el proceso de aprendizaje.<br>Facilita la visualización de conceptos abstractos y complejos.                         |
| 6  | La realidad aumentada y realidad virtual en la enseñanza matemática: educación inclusiva y rendimiento académico                              | (Martínez, 2024)               | Experiencias de aprendizaje personalizadas y atractivas.<br>Resolución práctica y teórica de problemas matemáticos mediante tecnologías inmersivas.   | Matemáticas<br>Educación inclusiva<br>Formación de futuros docentes                         | Desafíos técnicos y logísticos en la implementación de RA/RV.<br>Necesidad de capacitación para educadores en el uso de estas tecnologías   | Mejora significativa en el rendimiento académico de los estudiantes.<br>Fomento de experiencias de aprendizaje más inclusivas y accesibles                                  |



|    |   |   |  |  |   |  |
|----|---|---|--|--|---|--|
| 7  | Realidad Aumentada Móvil: Una estrategia pedagógica en el ámbito universitario                  | (Del Zulia et al., 2020)  | Visualización y manipulación de objetos virtuales tridimensionales. Fomento del aprendizaje del razonamiento espacial en estudiantes.  | Ingeniería industrial<br>Diseño industrial<br>Dibujo técnico<br>Educación universitaria                      | Dificultades técnicas en la implementación de la realidad aumentada.<br>Necesidad de formación docente en el uso de nuevas tecnologías  | Aumento de la motivación y el compromiso de los estudiantes.<br>Mejora en la comprensión de conceptos complejos a través de la visualización.  |
| 8  | La innovación en el aula universitaria a través de la realidad aumentada.                       | (Cabero Almenara, Vázquez Cano, Villota Oyarvide, et al., 2021) | Aplicaciones móviles: QuiverVisión y HP Reveal.<br>Simulaciones interactivas y visualizaciones en 3D.  | Educación Superior.<br>Ciencias de la Educación.<br>Tecnologías Emergentes.                                  | Falta de formación docente.<br>Limitaciones en la infraestructura tecnológica de algunos centros educativos.  | Mejora en las habilidades cognitivas y competencias en los estudiantes.<br>Aumento de motivación estudiantil en el aprendizaje educativo.  |
| 9  | La realidad virtual en escuelas hospitalarias.  | (Mente-clara et al., 2022)                                      | Simulaciones de eventos (por ejemplo, eventos sísmicos, exploración del interior del cuerpo humano, sistema solar).<br>Videos 360 que permiten a los estudiantes experimentar espacios, lugares, historias y pinturas. | Ciencias Naturales.<br>Ciencias Sociales.<br>Matemáticas.<br>Educación. Artes.<br>Tecnología.<br>Literatura. | Necesidad de capacitación específica para docentes.<br>Restricciones en el uso de la realidad virtual para ciertos estudiantes con condiciones médicas (por ejemplo, problemas de cuello, hipoacusia, convulsiones).<br>Requerimiento de equipamiento adecuado. | Permite a los estudiantes en hospitales acceder a experiencias educativas que de otro modo serían imposibles.<br>Aumenta el compromiso y la motivación de los estudiantes.                           |
| 10 | Nuevas competencias digitales en estudiantes potenciadas con el uso de Realidad Aumentada.      | (González Vidal et al., 2021)                                   | Experiencias didácticas que incluyen la visualización de contenidos educativos a través de dispositivos móviles.   | Educación (Pedagogía)<br>Tecnología educativa  | Falta de competencias digitales en algunos estudiantes, lo que dificulta el desarrollo de tareas.   | Mejora en la comprensión de conceptos complejos a través de experiencias interactivas.<br>Aumento de la motivación y satisfacción de los estudiantes al concluir actividades con RA.                 |
| 11 | Realidad aumentada para el fortalecimiento del pensamiento matemático geométrico espacial       | (Barragán Sánchez et al., 2024)                                 | Implementación de una estrategia didáctica apoyada por una aplicación de realidad aumentada para mejorar el pensamiento matemático geométrico espacial.  | Matemáticas<br>Geometría   | Dificultades para interpretar figuras geométricas por parte de los estudiantes.<br>Posibles barreras en la adaptación curricular para integrar la realidad aumentada.   | Aumento en el rendimiento académico de los estudiantes.  |
| 12 | Incidencia de la realidad aumentada en los procesos de aprendizaje de las funciones matemáticas | (Martínez et al., 2021)   | Aplicaciones de realidad aumentada para visualizar gráficamente conceptos teóricos de funciones matemáticas en 3D.   | Matemáticas.<br>Educación básica.  | Falta de capacitación docente para el uso efectivo de la realidad aumentada en el aula.<br>Limitaciones técnicas y de acceso a dispositivos móviles adecuados para todos los estudiantes.   | Mejora significativa en el aprendizaje de funciones matemáticas, con el grupo experimental superando al grupo de control en el post-test.<br>Aumento en la motivación de aprendizaje en matemáticas. |

|    |  |   |   |   |  |
|----|--|---|---|---|--|
|    |  |   |   |   |  |
| 13 | Realidad aumentada en educación universitaria  | (Cárdenas Castellanos et al., 2024)                             | Visualización de eventos históricos y personajes a través de modelos 3D interactivos.   | Historia<br>Educación básica  | Falta de recursos tecnológicos en algunas escuelas, lo que limita el acceso a la realidad aumentada.<br>Necesidad de formación docente para implementar efectivamente la realidad aumentada en el aula.<br><br>Aumento del interés y la participación de los estudiantes en el aprendizaje de la historia.<br>Mejora en la comprensión de eventos y conceptos históricos a través de experiencias inmersivas.          |
| 14 | Los efectos de la realidad virtual y la realidad aumentada en las actitudes hacia la ciencia en alumnos mexicanos de nivel primaria.               | (Salgado Reveles, 2023)   | Uso de aplicaciones como Chromville, Quiver, Merge Explorer y Expediciones para explorar conceptos científicos de manera interactiva.   | Ciencias Naturales.<br>Educación.   | Falta de infraestructura tecnológica, como la falta de dispositivos adecuados y acceso a internet.<br>Necesidad de capacitación al docente en realidad aumentada.<br><br>Buen desarrollo de actitudes hacia la ciencia, aumentando el interés de aprendizaje por parte de los estudiantes.<br>Mejora en la comprensión de conceptos científicos complejos a través de experiencias inmersivas y prácticas.             |
| 15 | Aplicación de Realidad Aumentada para Laboratorios de Química.   | (Hernández et al., 2021)  | Visualización de órganos y sistemas del cuerpo humano en 3D a través de dispositivos móviles.   | Biología<br>Ciencias de la Salud<br>Educación   | Alta dependencia de la tecnología y la necesidad de dispositivos móviles compatibles con las aplicaciones de RA.<br>Dificultades en la adaptación curricular para integrar la realidad aumentada en el aula de biología.<br><br>Aumento del interés y motivación estudiantil en el estudio de la anatomía humana.<br>Mejora en la comprensión de conceptos complejos a través de experiencias visuales e interactivas. |
| 16 | La innovación en el aula universitaria a través de la realidad aumentada. Análisis desde la perspectiva del estudiantado español y latinoamericano | (Cabero-Almenara, Vázquez-Cano, Villota-Oyarvide, et al., 2021) | Visualización de reacciones químicas y estructuras moleculares en 3D a través de dispositivos móviles.<br>Experimentos virtuales que permiten a los estudiantes interactuar con sustancias químicas.  | Química<br>Ciencias naturales   | Necesidad de dispositivos móviles adecuados y acceso a internet en el aula.<br>Falta de formación específica para docentes en el uso de herramientas de realidad aumentada.<br><br>Aumento de motivación estudiantil en Química.<br>Mejora en la comprensión de conceptos abstractos.  |
| 17 | Augmented Reality Application in Classroom: An Immersive Taxonomy  | (Kaviyaraj & Uma, 2022)   | Ánalisis de tecnologías de realidad aumentada en educación STEM.  | Matemáticas<br>Ciencias naturales.  | Confusión entre educadores, investigadores y desarrolladores.<br>Poco estudiado en educación STEM.<br><br>Mejora del aprendizaje y la motivación estudiantil.  |
| 18 | Augmented and virtual reality technologies in education  | (Muzyleva et al., 2021)   | Implementación de AR para la enseñanza de conceptos científicos y matemáticos.<br>Uso de RA en la formación profesional.<br>Creación de experiencias de aprendizaje interactivas en museos y exposiciones.<br>Desarrollo de aplicaciones educativas en el aprendizaje de idiomas. | Educación en ciencias.<br>Matemáticas.<br>Formación técnica y profesional<br>Educación artística<br>Idiomas | Problemas técnicos en la implementación de RA en entornos educativos.<br>Costos asociados con el desarrollo y mantenimiento de aplicaciones RA.<br>Necesidad de capacitación docente.<br><br>Mejora en la comprensión de conceptos complejos a través de visualizaciones interactivas.<br>Aumento de participación estudiantil<br>Colaboración de trabajo en equipo por parte de los estudiantes.                      |
| 19 | A Survey on Future of Augmented Reality with AI in Education   | (Kaviyaraj & Uma, 2021)   | Experiencias de aprendizajes interactivas.<br>Simulaciones en tiempo real.  | Educación   | Falta de investigación<br>Falta de experiencia en el personal docente.<br><br>Mejora en los educadores para las experiencias de aprendizaje mixtas.<br>Aumento en entornos de aprendizaje interactivo.   |



|    |   |   |  |  |   |
|----|---|---|--|--|---|
| 20 | Educational Advancements in the Field of Augmented Reality and Virtual Reality (Bheda et al., 2021)   | Uso de head-mounted displays y dispositivos móviles en educación.   | Educación en salud. Turismo. Marketing.  | Falta de una plataforma flexible para la creación de contenido con realidad aumentada. Desafíos encontrados seguridad y accesibilidad.   | Mayor efectividad y estimulación en el aprendizaje. Uso de AR/VR en la pandemia COVID-19.   |
| 21 | How Suitable is for Learners an Autonomous, Interactive and Dynamic Learning Model? (Real-Fernandez et al., 2021)   | Implementación de un sistema de aprendizaje autónomo e interactivo. Monitoreo del aprendizaje en tiempo real.   | Aplica en cualquier área educativa, está diseñado para personalizar la enseñanza según el ritmo y preferencias del estudiante. | Falta de herramientas diseñadas para la creación de contenido dinámico. Los sistemas educativos tradicionales aún no están totalmente preparados para personalizar el aprendizaje. | Mayor autonomía del estudiante, ajusta el ritmo del aprendizaje según sus necesidades. Interacción y motivación en los estudiantes para su proceso educativo.       |
| 22 | Augmented Reality and Virtual Reality for Learning: An Examination Using an Extended Technology Acceptance Model (Jang et al., 2021)                            | Simulaciones interactivas. Actividades de laboratorio con RA. Exploración de entornos virtuales.  | Física Biología Ciencias ambientales   | Falta de la capacitación docente. Barreras tecnológicas y de infraestructura. Resistencia al cambio por parte del personal docente.  | Mayor compromiso y motivación de los estudiantes. Aprendizaje inmersivo y práctico. Reducción de la ansiedad en pruebas y exámenes a los estudiantes.               |
| 23 | Applications of augmented reality (AR) in chemical engineering education: Virtual laboratory work demonstration to digital twin development (Zhou et al., 2024) | Visualización de patrones de flujo mediante experiencias de RA. Desarrollo de gemelos digitales con una interfaz de RA.                               | Ingeniería química Educación en ingeniería Procesos químicos   | Requerimientos técnicos y de procesamiento para experiencias de RA. Necesidad de habilidades específicas que permitan crear experiencias de RA.                                    | Mejora en la comprensión de conceptos complejos a través de visualizaciones interactivas. Reducción de pérdidas financieras.  |
| 24 | Augmented reality for chemical engineering education (Rebelo et al., 2024)  | Simulaciones interactivas. Representación de conceptos complejos en 3D.   | Ingeniería química. Química molecular. Operaciones unitarias. Educación técnica y profesional.                                 | Accesibilidad y costos de hardware para implementar RA. Dificultades en la integración de RA con simulaciones complejas como CFD. Necesidad de formación para el personal docente. | Aumento del compromiso y la motivación estudiantil. Mejora en la comprensión de áreas del currículo educativo complicadas a través de visualizaciones interactivas. |
| 25 | Molecule World: Enhancing chemistry education through web-based augmented reality using Assembler (Handoyo et al., 2024)  | Entorno de aprendizaje interactivo. Visualización en 3D de moléculas y compuestos. Interacción con estructuras moleculares a través de tecnología RA. | Química. Ciencias naturales Tecnología educativa.  | Dependencia en los dispositivos móviles. Problemas en la conectividad de internet. Necesidad de formación para estudiantes y educadores en el uso de RA.                           | Aumento del interés y la motivación en el aprendizaje de química. Mejor comprensión de conceptos moleculares complejos. Fomento de un aprendizaje más interactivo.  |

|    |   |                                |   |   |  |  |
|----|---|--------------------------------|---|---|--|--|
|    | Exploring the Impact of Virtual Reality and Augmented Reality Technologies in Sustainability Education on Green Energy and Sustainability Behavioral Change: A Qualitative Analysis | (Negi, 2024)                   | Simulaciones interactivas. Visualización en eternos virtuales de sistemas energéticos renovables.   | Energías renovables. Educación en sostenibilidad. Ciencias ambientales.                       | Costos asociados con la implementación. Problemas de accesibilidad y disponibilidad de dispositivos adecuados. Necesidad de formación continua en el personal docente.               | Aumento en comportamientos sostenibles entre estudiantes y profesionales.  |
| 26 | Introducing a general-purpose augmented reality platform for the use in engineering education   | (Grodotzki et al., 2023)       | Visualización en 3D de conceptos abstractos y complejos. Simulaciones interactivas en laboratorios. Juegos educativos basados en problemas.                             | Educación en ciencias. Matemáticas. Artes. Historia. Ingeniería                               | Costos de implementación y mantenimiento de tecnologías RA. Falta de infraestructura tecnológica adecuada. Necesidad de capacitación para el personal docente en el uso de la RA.    | Aumento del compromiso y la motivación estudiantil. Mejor comprensión de conceptos a través de las visualizaciones interactivas.   |
| 27 | Analyzing augmented reality (AR) and virtual reality (VR) recent development in education   | (Abdullah et al., 2023)        | Creación de entornos de aprendizaje inmersivos mediante RA. Visualización de datos en tiempo real. Simulaciones de escenarios del mundo real.                           | Educación superior. Ciencias aplicadas. Educación en ingeniería. Tecnología de la información | Costos elevados de hardware y software necesarios para RA. Desigualdades en el acceso a la tecnología entre estudiantes. Necesidad de formación y soporte técnico para los docentes. | Aumento del interés y la motivación de los estudiantes. Mejores resultados académicos.   |
| 28 | Using innovative technology tools in organic chemistry education: bibliometric analysis   | (Pabuçcu-Akış, 2024)           | Visualización de estructuras moleculares 3D.  | Química orgánica.   | Acceso y recursos. Adaptación curricular.  | Mejoras en el aprendizaje y comprensión. Aprendizaje interactivo y autónomo.   |
| 29 | Augmented reality in developing students' understanding of chemistry triplet: a systematic literature review.   | (Ribić & Devetak, 2024)        | Visualización de estructuras y fenómenos químicos. Simulación de experimentos de laboratorio. Aprendizaje inmersivo e interactivo.                                      | Química   | Problemas técnicos. Falta de integración pedagógica. Accesibilidad de dispositivos.  | Actitudes y motivación positiva. Mejoras en el aprendizaje. Valor pedagógico percibido por los docentes.   |
| 30 | Exploring the knowledge and awareness on applications of virtual reality and augmented reality technology among dental healthcare professionals—a crosssectional survey.            | (Masood et al., 2024)          | Educación dental simulada. Planificación quirúrgica y procedimientos clínicos.  | Odontología general y especializada. Cirugía oral y maxilofacial.                             | Falta de conocimiento. Costos elevados.  | Interés y disposición a adoptar la tecnología. Mejora en la educación clínica.   |
| 31 | Enriching a Traditional Learning Activity in Preschool through Augmented Reality: Children's and Teachers' Views.   | (Bacca-Acosta et al., 2023)    | Integración de realidad aumentada en actividades tradicionales de enseñanza en preescolar.  | Educación preescolar.   | Dificultades técnicas. Insuficiente disponibilidad de dispositivos.  | Mayor motivación y compromiso de los niños en el aprendizaje. Apoyo en la comprensión de conceptos abstractos, como el sistema solar.                                    |
| 32 | The Impact of Augmented Reality (AR) on the Academic Performance of High School Students.   | (Amores-Valencia et al., 2023) | Implementación de una aplicación de realidad aumentada (ComputAR). Comparación entre un grupo que utilizó AR y otro con metodología tradicional basada en diapositivas. | Tecnología e informática. Metodologías educativas.  | Possible dificultad de acceso a dispositivos móviles compatibles con la aplicación. Necesidad de formación previa de los docentes para una correcta implementación.                  | Mejora significativa del rendimiento académico de los estudiantes que usaron AR en comparación con los métodos tradicionales. Mayor motivación y participación en clase. |



|    |  |                                |  |   |  |   |
|----|--|--------------------------------|--|---|--|---|
| 34 | Enhanced Visualisation of Normal Anatomy with Potential Use of Augmented Reality Superimposed on Three-Dimensional Printed Models.   | (Geerlings-Batt et al., 2022)  | Desarrollo de herramientas de enseñanza de anatomía mediante realidad aumentada y modelos impresos en 3D. Uso de RA en combinación con modelos físicos para la comprensión de estructuras anatómicas.        | Medicina y ciencias de la salud. Educación en tecnología médica.    | Alto costo y accesibilidad limitada de las herramientas. Dificultades en el reconocimiento de objetos y superposición de RA en modelos físicos.                            | Mayor interactividad y realismo en la enseñanza de anatomía. Posibilidad de manipular modelos físicos y virtuales para mejorar la comprensión espacial.                         |
| 35 | Effects of learner control design in an AR-based exhibit on visitors' museum learning.   | (Lin et al., 2022)             | Implementación de realidad aumentada en exhibiciones de museos para mejorar la experiencia de aprendizaje de los visitantes.   | Educación en museos.  | Distracción de los visitantes debido a la falta de control sobre el contenido en algunos niveles de interacción. Necesidad de capacitación previa para los visitantes.     | Mayor compromiso y motivación de los visitantes al permitirles interactuar activamente con el contenido. Mejora en la adquisición de conocimientos mediante herramientas de RA. |
| 36 | AR4FSM: Mobile Augmented Reality Application in Engineering Education for Finite-State Machine Understanding.                        | (Nadeem et al., 2022)          | Visualización interactiva de máquinas de estados finitos (FSM) mediante realidad aumentada. Uso de multimedia (texto, imágenes, sonido y animaciones) para facilitar el aprendizaje.                         | Ingeniería de sistemas computacionales. Ciencias de la computación. | Dependencia en los dispositivos móviles compatibles con realidad aumentada. Necesidad de adaptación curricular.  | Mayor compromiso y motivación del estudiante. Facilidad en la comprensión de conceptos complejos mediante interacción inmersiva.  |
| 37 | A Review on Augmented Reality Authoring Toolkits for Education.  | (Dengel et al., 2022)          | Evaluación de herramientas de autoría de realidad aumentada.   | Tecnología educativa. Desarrollo de software educativo.             | Falta de evaluación empírica en muchas herramientas de RA. Pocas opciones para el personal docente que no tiene conocimientos técnicos.                                    | Identificación de cinco herramientas accesibles. Propuesta de criterios basados en evidencia para el desarrollo de futuras herramientas de RA.                                  |
| 38 | Designing augmented reality for makerspaces: Guidelines, lessons and mitigation strategies from 5+ years of AR educational projects. | (Radu & Schneider, 2023)       | Aplicación de realidad aumentada en espacios de creación (makerspaces) para mejorar el aprendizaje y la colaboración. Desarrollo de cinco proyectos de investigación que integran RA en entornos educativos. | Educación STEM.   | Sobrecarga cognitiva debido a la cantidad de información visual presentada en RA. Necesidad de medir la experiencia según el nivel de experiencia que tengan los usuarios. | Mejora en la comprensión de conceptos complejos mediante la visualización. Aumento en la motivación y participación de los estudiantes.   |
| 39 | The potential of mobile augmented reality as a didactic and pedagogical source in learning geometry 3D.                              | (Yaniawati et al., 2023)       | Integración de realidad aumentada en la enseñanza de geometría tridimensional.   | Matemáticas. Tecnología educativa.                                  | Dependencia de dispositivos móviles compatibles con la realidad aumentada. Necesidad de formación previa de docentes y estudiantes para una correcta integración.          | Mejora en la comprensión de conceptos geométricos mediante visualización interactiva. Mejor motivación y compromiso en el aprendizaje.  |
| 40 | The Influence of Augmented Reality (AR) on the Motivation of High School Students.   | (Amores-Valencia et al., 2023) | Implementación de una aplicación de realidad aumentada (ComputAR).   | Tecnología e informática. Psicología educativa.                     | Necesidad de formación previa a docentes y estudiantes para una correcta implementación en el ámbito educativo.  | Mayor interés en el proceso de aprendizaje. Incremento de la motivación estudiantil.  |

|    |   |                            |  |   |   |  |
|----|---|----------------------------|--|---|---|--|
| 41 | A Systematic Review of the Use and Effect of Virtual Reality, Augmented Reality and Mixed Reality in Physical Education.                  | (Pérez-Muñoz et al., 2024) | Aplicación de realidad virtual, aumentada y mixta en educación física.   | Educación física. Salud y bienestar. Inclusión educativa      | Necesidad de dispositivos tecnológicos avanzados. Posibles dificultades en la integración curricular.   | Motivación y compromiso de los estudiantes. Facilidad en el aprendizaje de habilidades motoras mediante experiencias inmersivas.   |
| 42 | The Potential of a Mobile Augmented Reality Game in Education for Sustainability: Report and Analysis of an Activity with the EduCITY App | (Rodrigues & Pombo, 2024)  | Implementación de un juego de realidad aumentada en la aplicación EduCITY para promover la educación sobre sostenibilidad.   | Educación ambiental. Ciencias Sociales. Tecnología educativa. | Dependencia en los dispositivos móviles compatibles con RA. Necesidad de una correcta formación previa para los docentes y estudiantes en el uso de la aplicación.            | Mayor motivación y compromiso de los estudiantes en su aprendizaje. Facilita la comprensión de conceptos ambientales mediante experiencias interactivas.   |
| 43 | The impact of augmented reality on education: a bibliometric exploration.   | (Singh et al., 2024)       | Visualización tridimensional de estructuras como huesos y órganos.   | Física. Química. Anatomía. Matemáticas. Arquitectura.         | Falta de formación docente. Accesibilidad y costos.   | Motivación de los estudiantes, facilitando retención y comprensión de información.   |
| 44 | Virtual reality and augmented reality in medical education: an umbrella review.   | (Tene et al., 2024)        | Evaluación del impacto de la realidad virtual y aumentada en la educación médica. Comparación entre métodos tradicionales y el uso de tecnologías inmersivas en la formación de profesionales de la salud. | Educación médica.   | Costos elevados en la implementación de las tecnologías inmersivas. Necesidad de capacitación en el personal docente para la correcta integración de RA y RV en el currículo. | Mejora en la retención del conocimiento y el desarrollo de sus habilidades a través de las prácticas. Mayor compromiso y motivación de los estudiantes en su aprendizaje.                            |
| 45 | Enhancing digital literacy in primary education through augmented reality.  | (Nevrelova et al., 2024)   | Implementación de la aplicación Quiver en la enseñanza de geometría en educación primaria. Uso de RA para mejorar la comprensión de sólidos platónicos y conceptos matemáticos abstractos.                 | Matemáticas. Educación primaria.                              | Dependencia en los dispositivos móviles compatibles con realidad aumentada. Necesidad de formación previa para docentes y estudiantes en el uso de la aplicación con RA       | Mayor motivación y compromiso de los estudiantes con el uso de la aplicación. Facilidad en la comprensión de conceptos geométricos mediante visualización interactiva.                               |
| 46 | Grand Challenges for Augmented Reality.   | (Billingham, 2021)         | Aplicación de RA en educación para el uso remoto y simulaciones médicas.   | Ciencias de la computación. Ingeniería y tecnología.          | Alto costo y accesibilidad de los dispositivos avanzados que permiten integrar RA.  | Mejora en la interacción y la colaboración remota con la integración de RA. Mejora en la educación por sus experiencias inmersivas.  |
| 47 | Augmented reality-based learning for the comprehension of cardiac physiology in undergraduate biomedical students.                        | (González et al., 2020)    | Uso de una aplicación de RA en dispositivos móviles para un mejor aprendizaje con el ciclo cardíaco.   | Fisiología humana. Educación Biomédica.                       | Dependencia en los dispositivos móviles que son compatibles con la RA. Necesidad de una formación previa en el personal docente y estudiantes en el uso de esta tecnología.   | Alta mejora en la comprensión de conceptos anatómicos y fisiológicos del ciclo cardíaco. Aumento en el aprendizaje de los estudiantes para representar gráficamente procesos fisiológicos complejos. |
| 48 | ARPocketLab—A Mobile Augmented Reality System for Pedagogic Applications.   | (Nunes et al., 2024)       | Laboratorio virtual móvil ARPocketLab. Visualización de resultados y fenómenos físicos/ químicos.  | Física. Química.  | Diferencias de precisión entre métodos de seguimiento por marcador y por superficie. Limitaciones técnicas y de uso en los dispositivos.                                      | Alta usabilidad y retroalimentación positiva en el uso del laboratorio virtual. Interactividad y aprendizaje autónomo.   |



|    |  |                            |  |              |  |   |
|----|--|----------------------------|--|--------------|--|---|
| 49 | Augmented Reality Escape Classroom Game for Deep and Meaningful English Language Learning. | (Voreopoulos et al., 2024) | Juego de escape con RA para el aprendizaje del idioma inglés.<br>Narrativa inmersiva con RA y gamificación.  | Inglés.      | El 40% de los participantes enfrentaron dificultades técnicas con los dispositivos.<br>Falta de la capacitación docente. | El 90% de los docentes está de acuerdo en que el juego es muy motivador, divertido e inmersivo. Mejoras en las habilidades receptivas y productivas del idioma. |
| 50 | Augmented Reality in Primary Education: An Active Learning Approach in Mathematics.        | (Volioti et al., 2023)     | Aplicación educativa Cooking Math, son 9 juegos educativos con realidad aumentada que abordan temas matemáticos a través de recetas de cocina simuladas. | Matemáticas. | Estudiantes y docentes mencionaron la falta de ayuda dentro del sistema y que la interfaz puede ser confusa para niños.  | Fomenta la resolución de problemas, la exploración y la participación crítica a través de situaciones prácticas del día a día.                                  |

**Tabla 6.** Casos de uso encontrados en los artículos seleccionados.

Fuente: Los autores.

| Casos de uso encontrados                             | Frecuencia | %     | Artículos que los mencionan            |
|--|------------|-------|--|
| Visualización de conceptos en 3D                     | 10         | 20.0% | 12, 13, 15, 16, 24, 25, 27, 29, 34, 39 |
| Exploración de entornos virtuales                    | 6          | 12.0% | 7, 16, 22, 26, 34, 40                  |
| Simulaciones interactivas                            | 5          | 10.0% | 1, 22, 24, 26, 27                      |
| Juegos educativos                                    | 4          | 8.0%  | 27, 42, 49, 50                         |
| Experiencias inmersivas en educación STEM            | 4          | 8.0%  | 17, 25, 26, 38                         |
| RA aplicado en la enseñanza de química               | 2          | 4.0%  | 18, 25                                 |
| RA en la enseñanza de historia y patrimonio cultural | 2          | 4.0%  | 15, 20                                 |
| RA en educación técnica y profesional                | 2          | 4.0%  | 20, 26                                 |
| RA en Educación inclusiva y personalizada            | 2          | 4.0%  | 6, 18                                  |

**Tabla 7.** Áreas del currículo encontrados en los artículos seleccionados.

Fuente: Los autores.

| Áreas del currículo        | Frecuencia | %     | Artículos que las mencionan                    |
|----------------------------|------------|-------|--|
| Matemáticas                | 13         | 26.0% | 1, 5, 6, 9, 11, 12, 17, 18, 27, 39, 43, 45, 50 |
| Química                    | 10         | 20.0% | 4, 5, 16, 23, 24, 25, 29, 30, 43, 48           |
| Tecnología Educativa       | 6          | 12.0% | 10, 25, 26, 37, 39, 42                         |
| Ciencias de la salud       | 6          | 12.0% | 15, 20, 31, 34, 41, 44                         |
| Ciencias Naturales         | 5          | 10.0% | 4, 9, 16, 17, 25                               |
| Física                     | 4          | 8.0%  | 5, 22, 43, 48                                  |
| Ciencias                   | 3          | 6.0%  | 1, 3, 5  |
| Biología                   | 3          | 6.0%  | 4, 15, 22                                      |
| Ciencias Ambientales       | 3          | 6.0%  | 4, 22, 26                                      |
| Idiomas                    | 2          | 4.0%  | 1, 18, 49                                      |
| Ciencias de la computación | 2          | 4.0%  | 36, 46   |

la salud se destacan con un 12% cada una, reflejando el uso de RA en entornos de formación especializada, ya sea en la mejora de estrategias pedagógicas o en la simulación de procedimientos clínicos y quirúrgicos que enriquecen la formación en el ámbito de la salud. Se identifican también otras áreas con aplicaciones significativas de la RA, como Ciencias naturales, que representa

el 10% y Física con un 8%, donde esta herramienta ha sido utilizada principalmente para ilustrar fenómenos abstractos y dinámicos. También se puede detectar una adopción creciente de la RA en disciplinas como Ciencias, Biología y Ciencias Ambientales. Por otra parte, en áreas como Idiomas y Ciencias de la Computación, cuya representación alcanza el 4%, aún existe

**Tabla 8.** Limitaciones encontradas en los artículos seleccionados.

Fuente: Los autores.

| Limitaciones  | Frecuencia | %      | Artículos que los mencionan  |
|---|------------|--------|--|
| Falta de formación docente  | 30         | 60.00% | 1, 2, 3, 5, 6, 7, 8, 9, 12, 13, 14, 16, 17, 18, 19, 22, 24, 26, 27, 28, 33, 37, 39, 40, 42, 43, 44, 45, 47, 49 |
| Problemas técnicos  | 29         | 58.00% | 1, 3, 4, 5, 6, 7, 13, 14, 15, 18, 20, 23, 24, 26, 27, 28, 30, 32, 34, 35, 36, 38, 39, 41, 43, 44, 46, 48, 50   |
| Acceso limitado a dispositivos móviles o infraestructura tecnológica      | 27         | 54.00% | 2, 3, 4, 8, 10, 12, 13, 14, 15, 16, 23, 24, 25, 26, 27, 28, 30, 32, 33, 34, 39, 42, 43, 45, 46, 48, 50         |
| Costos elevados   | 23         | 46.00% | 1, 5, 6, 9, 14, 18, 20, 24, 26, 27, 28, 31, 32, 34, 38, 41, 43, 44, 45, 47, 48, 49, 50                         |
| Adaptación curricular o integración pedagógica difícil                    | 17         | 34.00% | 3, 6, 11, 15, 21, 29, 30, 36, 38, 40, 41, 42, 44, 46, 47, 48, 50   |
| Falta de experiencia previa (docentes o estudiantes)                      | 15         | 30.00% | 1, 5, 10, 17, 19, 24, 25, 31, 37, 38, 39, 40, 47, 48, 50   |
| Conectividad o acceso a internet limitado                                 | 9          | 18.00% | 14, 16, 25, 26, 30, 34, 42, 48, 50   |
| Barreras de accesibilidad   | 6          | 12.00% | 6, 9, 20, 28, 34, 49   |
| Resistencia al cambio   | 4          | 8.00%  | 17, 19, 22, 47   |
| Falta de plataformas o herramientas adecuadas para creación de contenidos | 4          | 8.00%  | 20, 21, 37, 49   |
| Sobrecarga cognitiva  | 2          | 4.00%  | 35, 38   |
| Limitaciones médicas o de salud   | 1          | 2.00%  | 9  |

un amplio potencial para explorar e innovar con el uso de la RA.

De esta manera, la respuesta a la pregunta de investigación planteada es que la realidad aumentada ha tenido un mayor impacto en áreas del currículo educativo relacionadas con Matemáticas y Química, puesto que estas disciplinas requieren herramientas que faciliten la visualización, comprensión y asimilación de conceptos abstractos y el aprendizaje experimental.

### PI3: ¿Qué limitaciones enfrentan las aplicaciones de RA en dispositivos móviles en el ámbito educativo?

Los estudios analizados muestran que, a pesar su enorme crecimiento en el ámbito educativo, la RA también enfrenta desafíos que limitan su implementación y uso en el desarrollo de los contenidos curriculares. La tabla 8 muestra las principales limitaciones, organizadas ascendentemente por frecuencia.

Mediante el análisis de las limitaciones, se encontró que las principales se relacionan con la falta de formación docente y la necesidad de capacitación continua, obteniendo un 60% de frecuencia. Esto radica en que, si bien los docentes pueden tener una formación inicial, no todos tienen los conocimientos y habilidades para integrar de manera efectiva la RA en sus prácticas educativas; en consecuencia, la capacitación continua de los educadores es esencial para actualizar sus competencias tecnológicas y adaptarse a las herramientas vigentes. De no ser así, se limita el potencial de esta herramienta para transformar la enseñanza y el aprendizaje.

Le sigue los problemas técnicos como fallos en el funcionamiento, incompatibilidad de sistemas y la falta de soporte técnico adecuado, con una frecuencia del 58%, mientras que el acceso

limitado a dispositivos móviles o infraestructura tecnológica un 54%, demostrando que se requiere de dispositivos móviles avanzados y con buen procesamiento gráfico, logrando generar exclusión en los estudiantes sin acceso a equipos compatibles. Además, existen entornos educativos con menor desarrollo tecnológico que carecen de redes y plataformas robustas que permitan la implementación.

Además de los aspectos técnicos y de acceso, los costos elevados fueron mencionados en un 46% de los estudios seleccionados, debido a que estos costos pueden representar un desafío en las instituciones educativas con presupuestos limitados, donde las prioridades se centran en recursos básicos antes que en tecnologías emergentes.

Otra de las limitaciones encontradas e importante es la barrera de adaptación curricular o integración pedagógica con un 34%, la incorporación de tecnologías innovadoras como la RA puede requerir modificaciones en los contenidos, métodos de enseñanza y evaluación, por lo que puede generar resistencia tanto en los educadores como en las instituciones educativas.

En la misma línea, la falta de experiencia previa de docentes y estudiantes en el uso de RA obtuvo un 30% de frecuencia, lo cual indica que el éxito de su adopción depende también del nivel de alfabetización tecnológica de los actores educativos. La ausencia de competencias digitales o de conocimientos específicos sobre el funcionamiento y potencial de la RA puede generar resistencia o uso limitado.

El acceso deficiente a internet se mencionó en el 18% de los estudios como un obstáculo persistente, sobre todo en entornos



cuya infraestructura digital está poco desarrollada. Esta limitación afecta directamente a las aplicaciones de RA que requieren acceso en línea, actualización constante de contenido o sincronización entre dispositivos, puesto que restringe su uso pleno en el aula de clases o en entornos de aprendizaje híbridos.

Una de las limitaciones con menor frecuencia (12%) corresponde a las barreras de accesibilidad, que contempla dificultades para estudiantes con discapacidad, así como también la desigualdad en el acceso a la tecnología. De aquí nace, además, la necesidad de diseñar experiencias de RA más inclusivas y adaptadas a todos los estudiantes, independientemente de sus condiciones físicas o socioeconómicas. La resistencia al cambio y la falta de plataformas o herramientas adecuadas para la creación de contenidos obtuvieron una representatividad del 8%. Por una parte, la resistencia al cambio es una barrera de carácter más actitudinal o cultural, donde la adopción de la RA padece aún del escepticismo o rechazo de docentes y directivos tradicionalistas; por otra parte, está la falta de plataformas o herramientas adecuadas para la creación de contenidos en RA lo que es resultado de la carencia de soluciones accesibles y versátiles que permitan a los docentes diseñar experiencias personalizadas según los objetivos pedagógicos, ya que muchas

de las plataformas existentes requieren conocimientos técnicos avanzados o están orientadas a desarrolladores, limitando su uso por parte del profesorado y reduciendo la autonomía en la generación de materiales adaptados al contexto educativo.

En última instancia, se identificó que el 4% de los artículos presentaban sobrecarga cognitiva; es decir, sobrecarga de estímulos visuales o sensoriales, que pueden producir experiencias de RA no del todo positivas o beneficiosas, dificultando la concentración y comprensión de la información. A pesar de su valor porcentual, este hallazgo destaca la importancia de diseñar experiencias equilibradas, que promuevan y potencien el aprendizaje sin exceder la capacidad cognitiva del estudiante.

**PI4: ¿Qué beneficios se han identificado en el uso de esta tecnología para mejorar el aprendizaje en los últimos 5 años?**

Diversos estudios han señalado que la realidad aumentada no solo incrementa la motivación y el compromiso de los estudiantes, sino que también facilita la comprensión de conceptos complejos y promueve un aprendizaje más activo e interactivo.

La siguiente tabla 9 muestra los beneficios más frecuentes encontrados en los artículos seleccionados para la revisión. Los beneficios identificados revelan que los principales impactos

**Tabla 9.** Beneficios encontrados en los artículos seleccionados.

Fuente: Los autores.

| Beneficios   | Frecuencia | %     | Artículos que los mencionan   |
|--|------------|-------|---|
| Aumento de la motivación y compromiso                        | 35         | 70.0% | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 30, 32, 33, 35, 36, 38, 39, 40, 42 |
| Mejora en la comprensión de conceptos complejos o abstractos | 32         | 64.0% | 3, 4, 5, 7, 10, 11, 12, 13, 14, 15, 16, 17, 18, 23, 24, 25, 26, 27, 28, 29, 30, 32, 34, 35, 36, 38, 39, 41, 42, 43, 44, 50        |
| Interacción e inmersión en el aprendizaje                    | 17         | 34.0% | 10, 13, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 30, 35, 36, 41, 42  |
| Mejora del rendimiento académico                             | 10         | 20.0% | 2, 6, 11, 12, 18, 28, 30, 33, 41, 45  |
| Mejoras en habilidades cognitivas                            | 7          | 14.0% | 1, 8, 20, 21, 22, 31, 49  |
| Fomento del aprendizaje colaborativo y trabajo en equipo     | 5          | 10.0% | 3, 18, 19, 25, 48   |
| Valoración positiva por parte de docentes y estudiantes      | 4          | 8.0%  | 30, 31, 37, 43  |
| Apoyo a la enseñanza de anatomía o ciencias de la salud      | 4          | 8.0%  | 15, 34, 44, 47  |
| Conciencia ambiental y sostenibilidad                        | 3          | 6.0%  | 26, 42, 46  |
| Reducción de la ansiedad o estrés académico                  | 2          | 4.0%  | 22, 44  |
| Ahorro de tiempo o recursos en laboratorios                  | 2          | 4.0%  | 23, 49  |



positivos están relacionados con el compromiso, la comprensión de conceptos y la participación de los estudiantes. Es así que el beneficio más destacado es la mejora del compromiso y la motivación estudiantil (70%), hallazgo que evidencia que la RA genera experiencias más atractivas y dinámicas, fomentando la interacción y el interés del estudiante en el proceso de aprendizaje.

En segundo lugar, con una frecuencia del 64 %, se resalta la contribución de la RA en la mejora de la comprensión de conceptos complejos, toda vez que se puede presentar información de manera interactiva y visual, haciendo que los estudiantes asimilen estos contenidos con mayor claridad y efectividad en menos tiempo. La interacción e inmersión en el aprendizaje obtuvo un 34%, este beneficio destaca cómo la RA ofrece a los estudiantes la oportunidad de involucrarse activamente en entornos educativos más interactivos, estimulantes y motivadores. Además, se destaca la mejora del rendimiento académico con un 20%, lo cual evidencia que la RA puede ejercer una influencia favorable en los resultados del aprendizaje, al proporcionar una mayor comprensión de conceptos complejos, fortalecer el aprendizaje basado en lo visual y personalizar los contenidos según el ritmo individual del estudiante, particularmente en disciplinas como las matemáticas, las ciencias o la anatomía. Otro de los beneficios relevantes es las mejoras en habilidades cognitivas con un 14% de frecuencia, englobando el desarrollo de capacidades como la atención, la memoria espacial, la resolución de problemas y la toma de decisiones, a la par de fomentar la autogestión del aprendizaje.

Por otra parte, el aprendizaje colaborativo y el trabajo en equipo, con una frecuencia del 10%, evidencia que la realidad aumentada no se limita a experiencias individuales, sino que también opera como una herramienta eficaz para estimular dinámicas grupales.

Se identificaron también otros beneficios que, aunque con menor frecuencia, aportan un valor significativo al proceso educativo. Uno de ellos es la valoración positiva por parte de docentes y estudiantes, con un 8%, lo que demuestra una notable aceptación y disposición para la incorporación de la realidad aumentada (RA) en el entorno educativo.

También se destaca su utilidad en la enseñanza de anatomía y ciencias de la salud (8%), áreas en las que facilita la visualización detallada de estructuras complejas. Se identifican también el impulso de la conciencia ambiental y la sostenibilidad, con un 6%, así como la disminución de la ansiedad académica y el ahorro de tiempo y recursos en laboratorios, con un 4% cada uno.

Estos datos reflejan el gran potencial de la RA para mejorar tanto la experiencia como la gestión en el ámbito educativo.

### **Triangulación de la información**

Se encontró que los casos de uso más frecuentes, como visualización de conceptos 3D (20%) y la exploración de entornos virtuales (12%), tienen mayor impacto en matemáticas (26%) y química (18%). Esto evidencia la utilidad de la RA en áreas de currículo educativo que requieren representaciones visuales detalladas para una mejora de comprensión del contenido que imparten los docentes.

En cuanto a las limitaciones encontradas para la RA en la educación se destacan: la falta de formación docente (60%), barrera particularmente relevante, debido a que la falta de formación y capacitación docente impide que los educadores integren eficazmente la RA en sus prácticas pedagógicas; y, problemas técnicos (58%) o la ausencia de soporte especializado, que limita el uso de simulaciones interactivas y la exploración de entornos virtuales que requieren dispositivos y recursos tecnológicos para su implementación.

Estas barreras afectan la adopción de tecnologías inmersivas en la educación; sin embargo, pese a las limitaciones identificadas, el uso y aplicación de RA en el ámbito educativo tiene beneficios significativos como la mejora del compromiso y motivación estudiantil (70%) y la mejora en la comprensión de conceptos complejos (64%). La triangulación de la información se resume en la Tabla 10, donde se presentan los casos de uso, áreas del currículo, limitaciones y beneficios de la implementación de la realidad aumentada en la educación, con sus respectivos porcentajes para facilitar la interpretación de los resultados obtenidos.

**Tabla 10.** Triangulación de información.

Fuente: Los autores.

| Categoría           | Aspecto  | Porcentaje (%) | Comentarios   |
|---------------------|--|----------------|---|
| Casos de uso        | Visualización de conceptos en 3D                             | 20%            | Facilita la comprensión de conceptos complejos, mejorando la comprensión visual del contenido.                    |
|                     | Exploración de entornos virtuales                            | 12%            | Uso frecuente en la enseñanza, especialmente en ciencias naturales y matemáticas.                                 |
| Áreas del currículo | Matemáticas  | 26%            | Mejora en la visualización y entendimiento de problemas matemáticos.  |
|                     | Química  | 20%            | Alta efectividad en la enseñanza de conceptos científicos   |
| Limitaciones        | Falta de formación docente                                   | 60%            | Impide la integración efectiva de la RA en la práctica pedagógica.  |
|                     | Problemas técnicos   | 58%            | Limitaciones de compatibilidad, fallos y la ausencia de soporte técnico dificultan su implementación efectiva.    |
| Beneficios          | Aumento de la motivación y compromiso                        | 70%            | Aumento significativo en la participación de los estudiantes durante el desarrollo de las clases.                 |
|                     | Mejora en la comprensión de conceptos complejos o abstractos | 64%            | Su implementación ayuda a los estudiantes a entender los temas complejos de unas áreas de su currículo educativo. |



## Discusión

Los resultados obtenidos en esta revisión sistemática evidencian un impacto significativo del uso de la realidad aumentada (RA) en dispositivos móviles en el ámbito educativo, especialmente en disciplinas como Matemáticas y Química, donde la visualización de conceptos abstractos y la exploración de entornos virtuales son fundamentales para una comprensión más profunda.

Martínez (2024) argumenta que el uso de RA mejora la inclusión y el rendimiento académico en matemáticas al facilitar la comprensión de figuras geométricas y operaciones complejas mediante modelos interactivos, criterio que coincide con lo reportado por Hurtado et al., (2024), quienes señalaron que la RA permite a los niños aprender palabras, comprender historias y mejorar habilidades matemáticas, lo que convierte a la RA en un recurso eficaz para fomentar el aprendizaje significativo en estas disciplinas. De igual forma, en el área de Química, López et al., (2024) encontraron que el uso de RA en la enseñanza de química en educación media mejora el interés y la comprensión de los contenidos curriculares.

La aplicación de la RA en el entorno educativo también ofrece experiencias de aprendizaje personalizadas que se adaptan a diferentes estilos de aprendizaje y necesidades individuales, aspecto que resulta especialmente relevante para garantizar que todos los estudiantes tengan la oportunidad de beneficiarse de las innovaciones tecnológicas, eliminando así la brecha existente.

No obstante, a pesar de sus múltiples beneficios, su implementación enfrenta varios desafíos significativos, entre ellos, la falta de preparación y capacitación docente y los problemas técnicos, que siguen siendo obstáculos recurrentes puesto que esta tecnología requiere dispositivos de alta capacidad, como tabletas o smartphones con especificaciones técnicas adecuadas que asegurarle una experiencia sin interrupciones a los usuarios.

Este último aspecto, a su vez, representa una barrera económica importante para muchas instituciones educativas, especialmente aquellas con recursos limitados ya que, a diferencia de las instituciones con mejor financiamiento cuyas posibilidades de incorporar la RA son mayores, estas pueden verse limitadas en su acceso a esta tecnología innovadora, lo cual genera brechas digitales entre estudiantes de diferentes contextos socioeconómicos.

En concordancia con lo ya expuesto, Lancheros-Bohorquez y Vesga-Bravo (2024) enfatizan que, si bien la RA en dispositivos móviles representa una poderosa herramienta para transformar la enseñanza de disciplinas complejas como Matemáticas y Química, su implementación efectiva y exitosa depende también del soporte técnico que tenga disponible la institución y de la capacitación docente, ya que estos aspectos constituyen factores claves para

cerrar las brechas existentes en el acceso a la innovación educativa.

Es importante resaltar que, un personal docente capacitado en el uso de la RA permite integrar eficazmente actividades que aumentan la retención de contenidos, promueven la participación de los estudiantes y fomenta el aprendizaje interactivo para diversas áreas del currículo educativo sin aumentar la carga cognitiva.

Esto respalda lo señalado por Lucero (2024), donde afirma que la RA puede utilizarse para superponer información digital en el entorno real, proporcionando a los futuros maestros herramientas y recursos adicionales mientras enseñan.

Además, es fundamental un trabajo en conjunto y estructurado con los tomadores de decisiones que, a través de sus políticas se fortalezca la inversión en infraestructuras tecnológicas y programas de formación que faciliten la implementación.

## 4. Conclusiones

La metodología empleada, basada en el método analítico y la revisión sistemática de Kitchham, permitió identificar patrones y tendencias en la investigación actual, lo que facilitó una síntesis clara y estructurada de los beneficios, desafíos y aplicaciones de la realidad aumentada en dispositivos móviles en el ámbito de la educación.

Los resultados obtenidos revelan que el uso de la realidad aumentada en dispositivos móviles dentro del ámbito educativo tiene un impacto positivo significativo en la enseñanza-aprendizaje de las áreas del currículo educativo donde la visualización y la interacción juegan un papel fundamental. Matemáticas y Química son las que más se benefician de su aplicación, seguidas de Ingeniería y Educación Técnica.

Tras el análisis de los resultados, se identificó un patrón claro: mientras que el impacto positivo de la RA en el aprendizaje es evidente, su adopción a gran escala enfrenta desafíos considerables.

Se identificaron limitaciones como el soporte técnico y la falta de preparación docente, que representan los principales cuellos de botella para la efectiva integración de estas tecnologías innovadoras en el entorno educativo.

Estas limitaciones sugieren que es fundamental invertir en formación continua para docentes, mejorar la infraestructura tecnológica y reducir las barreras económicas.

Además, es necesario elaborar contenido accesible y adaptable a diferentes contextos educativos, permitiendo una integración curricular más efectiva.

Para investigaciones futuras, se recomienda realizar estudios

empíricos que analicen la efectividad de la realidad aumentada en áreas curriculares menos exploradas, como la enseñanza de idiomas, ciencias de la computación y pedagogía en general.

Así como también investigar y proponer modelos de formación docente y estrategias institucionales que faciliten la implementación de la RA en el entorno educativo, a fin de garantizar un uso más eficiente y sostenible de esta tecnología.

### Agradecimientos

Queremos expresar un sincero agradecimiento a cada uno de los autores que aportan a la investigación a través del desarrollo de artículos científicos que posteriormente pueden ser utilizados para una revisión sistemática como el presente trabajo.

Los conocimientos y hallazgos que comparten han permitido investigar aplicaciones y beneficios de la realidad aumentada en la educación.

Agradecemos el compromiso y esfuerzo investigativo para seguir avanzando en el campo de la educación tecnológica.

### Contribución de los autores

**Jeniffer Karina Alcívar Pazmiño:** Investigación, Redacción - borrador original del artículo. **Yimmy Salvador Loor Vera:** Redacción - revisión y edición del artículo. **Joffre Edgardo Panchana Flores:** Metodología.

### Conflictos de interés

Los autores no tienen conflictos de intereses.

### Referencias bibliográficas

- Abdullah, A.-A., Mohammed, J., Askar, G., & Ahmed, A.-A. (2023). Analyzing augmented reality (AR) and virtual reality (VR) recent development in education. *Social Sciences & Humanities Open*, 8(1), 100532. <https://doi.org/10.1016/J.SSAHO.2023.100532>
- Aguilar Acevedo, F., Flores Cruz, J. A., Pacheco Bautista, D., & Caldera Miguel, J. (2023). Perspectiva tecnopedagógica de la realidad aumentada en la educación. *Investigación y Ciencia: De La Universidad Autónoma de Aguascalientes*, 31(90), 1–13. <https://doi.org/10.33064/iycuaa2023904252>
- Ali, O., Abdelbaki, W., Shrestha, A., Elbasi, E., Alryalat, M. A. A., & Dwivedi, Y. K. (2022). A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. *Journal of Innovation & Knowledge*, 8(1). <https://doi.org/10.1016/J.JIK.2023.100333>
- Amores-Valencia, A., Burgos, D., & Branch-Bedoya, J. W. (2023). The Impact of Augmented Reality (AR) on the Academic Performance of High School Students. *Electronics*, 12(10), 2173. <https://doi.org/10.3390/ELECTRONICS12102173>
- López, J., Rangel, J., & Calao, S. (2024). Realidad aumentada en el aula de química: estrategia multimedia para mejorar la comprensión de compuestos en estudiantes de secundaria. *Acta Scientiae Informaticae*, 8(8), 4–10. <https://doi.org/10.21897/26192659.3621>
- Arellano Pimentel, J., Solar González, R., & Armería Zavala, L. (2024). Estrategias y recursos didácticos utilizados para aprender programación estructurada. Una revisión sistemática. *IE Revista de Investigación Educativa de La REDIECH*, 15, e1872. [https://doi.org/10.33010/ie\\_rie\\_rediech.v15i0.1872](https://doi.org/10.33010/ie_rie_rediech.v15i0.1872)
- Bacca-Acosta, J., Duque-Mendez, N. D., Rapti, S., Sapounidis, T., & Tselegkaridis, S. (2023). Enriching a traditional learning activity in preschool through augmented reality: Children's and teachers' views. *Information*, 14(10), 530. <https://doi.org/10.3390/INFO14100530>
- Barragán Sánchez, C. C., Diaz León, J. A., & Renteria Vera, J. A. (2024). Realidad aumentada para el fortalecimiento del pensamiento matemático geométrico espacial. *Panorama*, 18(34), 153–168. <https://doi.org/10.15765/CG8KKN05>
- Bheda, R., Bhimani, D., Dharamshi, F., Sheth, S., Menon, R., Somra, R., Bhasuru, R., Mahajan, C., Gajbhiye, S., & Toradmalle, D. (2021). Educational advancements in the field of augmented reality and virtual reality. Proceedings - International Conference on Communication, Information and Computing Technology, ICCICT 2021. <https://doi.org/10.1109/ICCICT50803.2021.9509941>
- Billinghurst, M. (2021). Grand challenges for augmented reality. *Frontiers in Virtual Reality*, 2, 578080. <https://doi.org/10.3389/FRVIR.2021.578080>
- Amores-Valencia, A., Burgos, D., & Branch-Bedoya, J. W. (2023). The Influence of Augmented Reality (AR) on the Motivation of High School Students. *Electronics*, 12(22), 4715. <https://doi.org/10.3390/electronics12224715>
- Cabascango Trávez, G. (2023). El uso de la realidad aumentada en la enseñanza de ciencias: Un enfoque integrador en educación secundaria. *Revista Científica Kosmos*, 2(1), 39–50. <https://doi.org/10.62943/RCK.V2N1.2023.43>
- Cabero Almenara, J., Vázquez Cano, E., Villota Oyarvide, W. R., & López Meneses, E. (2021). La innovación en el aula universitaria a través de la realidad aumentada. *Educare*, 25(3), 1–17. <https://doi.org/10.7440/res64.2018.03>
- Cabero-Almenara, J., Vázquez-Cano, E., Villota-Oyarvide, W. R., & López-Meneses, E. (2021). La innovación en el aula universitaria a través de la realidad aumentada. Análisis desde la perspectiva del estudiantado español y latinoamericano. *Revista Electrónica Educare*, 25(3), 1–17. <https://doi.org/10.15359/REE.25-3.1>
- Calderón Imbaquingo, Z. L., Sam Anlas, C. A., Hubel Solis, B.,



- & Huancollo Quispe, D. (2024). Integración de realidad aumentada en la enseñanza de lenguaje en la educación superior: transformando la experiencia de aprendizaje. *Reincisol*, 3(6), 3390–3414. [https://doi.org/10.59282/reincisol.V3\(6\)3390-3414](https://doi.org/10.59282/reincisol.V3(6)3390-3414)
- Cárdenas Castellanos, J., Camacho Vázquez, M. R., Adaile Benítez, N. T., & Pérez Méndez, J. (2024). Realidad aumentada en educación universitaria. *Emerging Trends in Education*, 7(13), 117–128. <https://doi.org/10.19136/ETIE.A7N13.6300>
- Carrión, J., & Serrano, V. (2021). Revisión sistemática de literatura: características y funcionamiento respecto a los modelos BERT y SQuAD. *CEDAMAZ: Revista del Centro de Estudio y Desarrollo de la Amazonía*, 11(1), 79–86.
- Del Zulia, U., Luis, V., & Arredondo, A. L. (2020). Realidad Aumentada Móvil: Una estrategia pedagógica en el ámbito universitario. *Revista Técnica de Ingeniería de la Universidad del Zulia*, 43(3), 142–149. <https://doi.org/10.22209/rt.v43n3a04>
- Dengel, A., Iqbal, M. Z., Grafe, S., & Mangina, E. (2022). A review on augmented reality authoring toolkits for education. *Frontiers in Virtual Reality*, 3. <https://doi.org/10.3389/FVRIR.2022.798032>
- Geerlings-Batt, J., Tillett, C., Gupta, A., & Sun, Z. (2022). Enhanced visualisation of normal anatomy with potential use of augmented reality superimposed on three-dimensional printed models. *Micromachines*, 13(10), 1701. <https://doi.org/10.3390/MI13101701>
- Gonzalez, A. A., Lizana, P. A., Pino, S., Miller, B. G., & Merino, C. (2020). Augmented reality-based learning for the comprehension of cardiac physiology in undergraduate biomedical students. *Advances in Physiology Education*, 44(3), 314–322. <https://doi.org/10.1152/ADVAN.00137.2019>
- González Pérez, A., & Cerezo Cortijo, I. (2020). Implicaciones pedagógicas de la realidad aumentada para la mejora de la enseñanza de las ciencias en primaria. *Revista Interuniversitaria de Investigación En Tecnología Educativa*, 1–16. <https://doi.org/10.6018/riite.444961>
- González Vidal, I., Cebreiro López, B., & Casal Otero, L. (2021). Nuevas competencias digitales en estudiantes potenciadas con el uso de Realidad Aumentada. Estudio Piloto. RIED. *Revista Iberoamericana de Educación a Distancia*, 24(1). <https://doi.org/10.5944/ried.24.1.27501>
- Grodotzki, J., Müller, B. T., & Tekkaya, A. E. (2023). Introducing a general-purpose augmented reality platform for the use in engineering education. *Advances in Industrial and Manufacturing Engineering*, 6, 100116. <https://doi.org/10.1016/J.AIME.2023.100116>
- Handoyo, K. J., Wisnuwardana, C. J., Austen, A., & Permana, F. (2024). Molecule World: Enhancing chemistry education through web-based augmented reality using Assemblr. *Procedia Computer Science*, 245(C), 1249–1258. <https://doi.org/10.1016/J.PROCS.2024.10.354>
- Hernández, D., Bottner, E., Cataldo, F., & Zaragoza, E. (2021). Aplicación de Realidad Aumentada para Laboratorios de Química. *Educación Química*, 32(3), 30–37. <https://doi.org/10.22201/FQ.18708404E.2021.3.68129>
- Hurtado-Mazeyra, A., Condori-Yucra, N., Ponce-Alvarez, E., Limaymanta, C. H., & Suárez-Guerrero, C. (2024). Uso didáctico de la Realidad Aumentada en la Educación Infantil: Una revisión sistemática. *Revista Complutense de Educación*, 35(3), 515–528. <https://doi.org/10.5209/rced.85815>
- Jang, J., Ko, Y., Shin, W. S., & Han, I. (2021). Augmented Reality and Virtual Reality for Learning: An Examination Using an Extended Technology Acceptance Model. *IEEE Access*, 9, 6798–6809. <https://doi.org/10.1109/ACCESS.2020.3048708>
- Kamińska, D., Zwoliński, G., Laska-Leśniewicz, A., Raposo, R., Vairinhos, M., Pereira, E., Urem, F., Ljubić Hinić, M., Haamer, R. E., & Anbarjafari, G. (2023). Augmented Reality: Current and New Trends in Education. *Electronics (Switzerland)*, 12(16). <https://doi.org/10.3390/electronics12163531>
- Kaviyaraj, R., & Uma, M. (2021). A Survey on Future of Augmented Reality with AI in Education. Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021, 47–52. <https://doi.org/10.1109/ICAIS50930.2021.9395838>
- Kaviyaraj, R., & Uma, M. (2022). Augmented Reality Application in Classroom: An Immersive Taxonomy. Proceedings - 4th International Conference on Smart Systems and Inventive Technology, ICSSIT 2022, 1221–1226. <https://doi.org/10.1109/ICSSIT53264.2022.9716325>
- Lancheros-Bohorquez, W. F., & Vesga-Bravo, G. J. (2024). Uso de la realidad aumentada, la realidad virtual y la inteligencia artificial en educación secundaria: una revisión sistemática. *Revista de Investigación, Desarrollo*

- e Innovación, 14(1), 95–110. <https://doi.org/10.19053/UPTC.20278306.V14.N1.2024.17537>
- Leal Aragón, L. (2020). Producción de recursos didácticos para el aula de matemáticas de Secundaria con realidad aumentada. *Innovación Educativa*, 30, 185–198. <https://doi.org/10.15304/ie.30.6905>
- Lin, W., Lo, W. T., & Yueh, H. P. (2022). Effects of learner control design in an AR-based exhibit on visitors' museum learning. *PLOS ONE*, 17(10), e0274826. <https://doi.org/10.1371/journal.pone.0274826>
- Lucero Baldevenites, E. V. (2024). Transformando la educación: IA y realidades aumentadas y virtual en la formación docente. *European Public & Social Innovation Review*, 9, 1–16. <https://doi.org/10.31637/epsir-2024-854>
- Márquez Díaz, J. E., & Morales Espinosa, L. A. (2020). Realidad aumentada como herramienta de apoyo al aprendizaje de las funciones algebraicas y trascendentes. *Revista Educación En Ingeniería*, 15(29), 34–41. <https://doi.org/10.26507/rei.v15n29.1037>
- Martínez, J. O. (2024). La realidad aumentada y realidad virtual en la enseñanza matemática: educación inclusiva y rendimiento académico. *EduTec. Revista Electrónica de Tecnología Educativa*, 88(88), 62–76. <https://doi.org/10.21556/edutec.2024.88.3133>
- Martínez, O. M., Mejía, E., Ramírez, W. R., & Rodríguez, T. D. (2021). Incidencia de la realidad aumentada en los procesos de aprendizaje de las funciones matemáticas. *Información Tecnológica*, 32(3), 3–14. <https://doi.org/10.4067/S0718-07642021000300003>
- Masood, Z., Qabool, H., Fida, M., & Sukhia, R. H. (2024). Exploring the knowledge and awareness on applications of virtual reality and augmented reality technology among dental healthcare professionals – a cross-sectional survey. *Journal of the Pakistan Medical Association*, 74(4), S10–S16. <https://doi.org/10.47391/JPMA.AKU-9S-03>
- Menteclara, F., Roma, A., & Cecilia, M. (2022). La realidad virtual en escuelas hospitalarias. *Revista Científica Arbitrada de La Fundación MenteClara*, 7, 1–24. <https://doi.org/10.32351/rca.v7.290>
- Moreno Fuentes, E., Hidalgo Navarrete, J., Burgos Bolós, C., & Blanca de la Paz, S. de la. (2021). Aprendizaje integrado y colaborativo de ciencias a través de la realidad aumentada en educación infantil. *Etic@net: Revista Científica Electrónica de Educación y Comunicación en la Sociedad del Conocimiento*, 21(1), 214–229. <https://doi.org/10.30827/eticanet.v21i1.17012>
- Muzyleva, I., Yazykova, L., Gorlach, A., & Gorlach, Y. (2021). Augmented and virtual reality technologies in education. Proceedings - 2021 1st International Conference on Technology Enhanced Learning in Higher Education, 1, 617–623. <https://doi.org/10.1109/WEEF-GEDC53299.2021.9657378>
- Education, TELE 2021, 99–103. <https://doi.org/10.1109/TELE52840.2021.9482568>
- Nadeem, M., Lal, M., Cen, J., & Sharshir, M. (2022). AR4FSM: Mobile Augmented Reality Application in Engineering Education for Finite-State Machine Understanding. *Education Sciences*, 12(8), 555. <https://doi.org/10.3390/educsci12080555>
- Negi, S. K. (2024). Exploring the Impact of Virtual Reality and Augmented Reality Technologies in Sustainability Education on Green Energy and Sustainability Behavioral Change: A Qualitative Analysis. *Procedia Computer Science*, 236, 550–557. <https://doi.org/10.1016/J.PROCS.2024.05.065>
- Nevrelova, N., Korenova, L., Lavicza, Z., Bruzkova, N., & Schmid, A. (2024). Enhancing digital literacy in primary education through augmented reality. *Frontiers in Education*, 9, 1390491. <https://doi.org/10.3389/feduc.2024.1390491>
- Nunes, M., Adão, T., Shahrabadi, S., Capela, A., Carneiro, D., Branco, P., Magalhães, L., Morais, R., & Peres, E. (2024). ARPocketLab—A Mobile Augmented Reality System for Pedagogic Applications. *Computers*, 13(6), 148. <https://doi.org/10.3390/computers13060148>
- Núñez Zavala, C. X., Isín Vilema, M. D., Jiménez Granizo, C. N., & Peñafiel Barros, G. O. (2023). Inteligencia artificial y realidad aumentada como herramientas innovadoras en el desarrollo de cuentos educativos. *Informática y Sistemas*, 7(2), 87–95. <https://doi.org/10.33936/isrtic.v7i2.6156>
- Pabuçcu-Akış, A. (2024). Using innovative technology tools in organic chemistry education: bibliometric analysis. *Chemistry Teacher International*, 7(1), 141–156. <https://doi.org/10.1515/cti-2024-0055>
- Pérez-Muñoz, S., Castaño Calle, R., Morales Campo, P. T., & Rodríguez-Cayetano, A. (2024). A Systematic Review of the Use and Effect of Virtual Reality, Augmented Reality and Mixed Reality in Physical Education. *Information*, 15(9), 582. <https://doi.org/10.3390/info15090582>
- Radu, I., & Schneider, B. (2023). Designing augmented reality for makerspaces: Guidelines, lessons and mitigation strategies from 5+ years of AR educational projects. *Computers & Education: X Reality*, 2, 100026. <https://doi.org/10.1016/j.cexr.2023.100026>
- Real-Fernandez, A., Molina-Carmona, R., & Llorens-Largo, F. (2021). How suitable is for learners an autonomous, interactive and dynamic learning model? Proceedings of 2021 World Engineering Education Forum/Global Engineering Deans Council, WEEF/GEDC 2021, 617–623. <https://doi.org/10.1109/WEEF-GEDC53299.2021.9657378>
- Rebelo, C. M., Deiró, G. F., Knuutila, H. K., Moreira, L. C. de



- S., & Nogueira, I. B. R. (2024). Augmented reality for chemical engineering education. *Education for Chemical Engineers*, 47, 30–44. <https://doi.org/10.1016/j.ece.2024.04.001>
- Ribič, L., & Devetak, I. (2024). Augmented reality in developing students' understanding of chemistry triplet: A systematic literature review. *Chemistry Teacher International*, 7(1), 157–172. <https://doi.org/10.1515/cti-2024-0060>
- Rivas Rebaque, B., Gértrudix Barrio, F., & Gértrudix-Barrio, M. (2021). Análisis sistemático sobre el uso de la realidad aumentada en Educación Infantil. Edutec. *Revista Electrónica de Tecnología Educativa*, 76, 53–73. <https://doi.org/10.21556/edutec.2021.76.2053>
- Rodrigues, R., & Pombo, L. (2024). The potential of a mobile augmented reality game in education for sustainability: Report and analysis of an activity with the EduCITY App. *Sustainability*, 16(21), 9357. <https://doi.org/10.3390/su16219357>
- Rodríguez López, M. (2022). Desarrollo del aprendizaje basado en proyectos con realidad aumentada en educación secundaria para mejorar rendimientos en el aula de música. *Artseduca*, 32, 135–146. <https://doi.org/10.6035/artseduca.6272>
- Sacoto, F., Guillen, V. M., & Hevia Artime, I. (2024). El aprendizaje del lenguaje, un motor de desarrollo: Revisión sistemática de literatura. *Praxis Pedagógica*, 24(36), 102–131. <https://doi.org/10.26620/uniminuto.praxis.24.36.2024.102-131>
- Salgado Reveles, M. A. (2023). Los efectos de la realidad virtual y la realidad aumentada en las actitudes hacia la ciencia en alumnos mexicanos de nivel primaria. *PAAKAT: Revista de Tecnología y Sociedad*, 13(25), 1–30. <https://doi.org/10.32870/pk.a13n25.804>
- Sánchez Hayman, L. H., Morales Tamayo, Y., & Trujillo Ronquillo, D. F. (2024). Revisión sistemática de literatura sobre la incidencia de la tecnología CNC en la Industria 4.0. *Revista Científica y Tecnológica UPSE*, 11(2), 145–155. <https://doi.org/10.26423/rectu.v11i2.801>
- Singh, S., Kaur, A., & Gulzar, Y. (2024). The impact of augmented reality on education: A bibliometric exploration. *Frontiers in Education*, 9, 1458695. <https://doi.org/10.3389/feduc.2024.1458695>
- Tebes, G., Becker, P., Peppino, D., & Olsina, L. (2019, April). Especificación del modelo de proceso para una revisión sistemática de literatura. *Proceedings of Curran Associates Publisher* 2019. <https://www.researchgate.net/publication/333855959>
- Tene, T., Vique López, D. F., Valverde Aguirre, P. E., Orna Puente, L. M., & Vacacela Gomez, C. (2024). Virtual reality and augmented reality in medical education: An umbrella review. *Frontiers in Digital Health*, 6, 1365345. <https://doi.org/10.3389/fdgh.2024.1365345>
- Urbina López, M., Endara Estévez, M., Toapanta Mendoza, A., Guaras Pinango, M., & Quinchiguango Jitala, J. (2024). El uso de realidad aumentada en la enseñanza de ciencias naturales en educación básica. *Revista Científica Retos de La Ciencia*, 1(4), 224–238. <https://doi.org/10.53877/rc.8.19e.202409.18>
- Urresta-Yépez, R. F. (2024). Realidad aumentada en educación superior y el análisis de sus beneficios y desafíos. *Horizon Nexus Journal*, 2(2), 57–70. <https://doi.org/10.70881/hnj/v2/n2/39>
- Vera Mora, G. R., Sanz, C., Baldassarri, S., & Coma, T. (2023). Entornos virtuales de enseñanza y aprendizaje gamificados a la luz del concepto de presencia: Revisión sistemática de literatura. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, 33, e3. <https://doi.org/10.24215/18509959.33.e3>
- Volioti, C., Orovas, C., Sapounidis, T., Trachanas, G., & Keramopoulos, E. (2023). Augmented reality in primary education: An active learning approach in mathematics. *Computers*, 12(10), 207. <https://doi.org/10.3390/computers12100207>
- Voreopoulou, A., Mystakidis, S., & Tsinakos, A. (2024). Augmented reality escape classroom game for deep and meaningful English language learning. *Computers*, 13(1), 24. <https://doi.org/10.3390/computers13010024>
- Yaniawati, P., Sudirman, Mellawaty, Indrawan, R., & Mubarika, M. P. (2023). The potential of mobile augmented reality as a didactic and pedagogical source in learning geometry 3D. *Journal of Technology and Science Education*, 13(1), 4–22. <https://doi.org/10.3926/jotse.1661>
- Zhou, Z., Oveissi, F., & Langrish, T. (2024). Applications of augmented reality (AR) in chemical engineering education: Virtual laboratory work demonstration to digital twin development. *Computers & Chemical Engineering*, 188, 108784. <https://doi.org/10.1016/j.compchemeng.2024.108784>



## Innovación Educativa con Inteligencia Artificial: Generación y Validación de Guiones para Monólogos Educativos en la Enseñanza Universitaria

## Education Innovation with Artificial Intelligence: Generation and Validation of Scripts for Educational Monologues in University Teaching

### Autores

\* **Christiam Xavier Núñez Zavala** 

✉ cnunez@unach.edu.ec

**Jorge Noé Silva Castillo** 

✉ jsilva@unach.edu.ec

**Cristhy Nataly Jimenez Granizo** 

✉ cjimenez@unach.edu.ec

**Hernán Ramiro Pailiacho Yucta** 

✉ hpailiacho@unach.edu.ec

### Resumen

Este estudio examina el potencial de la Inteligencia artificial generativa para crear guiones de monólogos educativos de tipo stand-up en contextos universitarios. La investigación comparó dos sistemas de IA generativa (Claude y ChatGPT) para generar guiones didácticos posteriormente validados por expertos en artes escénicas y pedagogía, siguiendo una rúbrica específica. Los guiones validados fueron interpretados por estudiantes de un semillero de investigación frente a 40 estudiantes de primer semestre de Pedagogía de la Informática, quienes evaluaron su efectividad mediante una encuesta que midió humor, relevancia del mensaje, claridad, interacción y reflexión en temas como innovación, TICs, pedagogía, programación, robótica e informática. Los resultados mostraron una preferencia significativa (70%) por los guiones generados por Claude, con diferencias notables en claridad (+0.8), conexión con el público (+0.9) y aprendizaje percibido en pedagogía (+1.3). El estudio evidencia el valor de la IA generativa como herramienta complementaria para la creación de material educativo innovador, facilitando nuevas aproximaciones metodológicas en la enseñanza universitaria.

**Palabras clave:** Inteligencia artificial educativa; monólogos educativos; stand-up pedagógico; innovación pedagógica; tecnología educativa.

Universidad Nacional de Chimborazo,  
Facultad de Ciencias de la Educación,  
Humanas y Tecnologías, Riobamba,  
Chimborazo, Ecuador.

\*Autor para correspondencia

### Abstract

This study examines the potential of generative artificial intelligence to create educational stand-up monologue scripts in university contexts. The research compared two generative AI systems (Claude and ChatGPT) to generate didactic scripts subsequently validated by experts in performing arts and pedagogy, following a specific rubric. The validated scripts were performed by students from a research seedbed in front of 40 first-semester students of Computer Science Pedagogy, who evaluated their effectiveness through a survey that measured humor, message relevance, clarity, interaction, and reflection on topics such as innovation, ICTs, pedagogy, programming, robotics, and computer science. The results showed a significant preference (70%) for scripts generated by Claude, with notable differences in clarity (+0.8), connection with the audience (+0.9), and perceived learning in pedagogy (+1.3). The study demonstrates the value of generative AI as a complementary tool for creating innovative educational material, facilitating new methodological approaches in university teaching.

**Keywords:** Educational artificial intelligence; educational monologues; pedagogical stand-up; pedagogical innovation; educational technology.

### Comó citar el artículo:

Núñez Zavala, C. X., Silva Castillo, J. N., Jimenez Granizo, C. N. & Pailiacho Yucta, H. P. (2025). Innovación Educativa con Inteligencia Artificial: Generación y Validación de Guiones para Monólogos Educativos en la Enseñanza Universitaria. *Informática y Sistemas*, 9(2), 126-139. <https://doi.org/10.33936/isrtic.v9i2.7471>

Enviado: 22/04/2025

Aceptado: 26/06/2025

Publicado: 08/07/2025



## 1. Introducción

La educación superior enfrenta el desafío constante de mantenerse pertinente y efectiva en un mundo digitalizado y en constante cambio. Las metodologías tradicionales de enseñanza, aunque valiosas, frecuentemente resultan insuficientes para captar y mantener la atención de estudiantes habituados a interacciones digitales constantes. La incorporación de estrategias didácticas innovadoras ha demostrado en los últimos años un incremento significativo del compromiso estudiantil y la retención del aprendizaje (Veletsianos, 2012).

El humor como herramienta pedagógica ha ganado reconocimiento por su capacidad para crear entornos de aprendizaje más receptivos, transformar la ansiedad estudiantil y aumentar la retención de información (Berk, 2007). Los monólogos educativos inspirados en el stand-up comedy representan una estrategia innovadora que combina entretenimiento y contenido académico, estableciendo conexiones afectivas y emocionales con el material de estudio. Sin embargo, la creación de estos recursos requiere habilidades específicas y tiempo considerable de desarrollo, limitando su implementación generalizada.

Paralelamente, la Inteligencia artificial generativa ha alcanzado capacidades avanzadas para generar contenido educativo personalizado y sistemático para diversos niveles de aprendizaje. Los sistemas como ChatGPT y Claude han demostrado impacto significativo en la generación automática de textos educativos relevantes, ampliando las oportunidades de apoyo en la elaboración de materiales didácticos (Brown et al., 2020).

### 1.1 Estado del Arte

#### Inteligencia artificial Generativa en Educación

La investigación reciente sobre IA generativa en contextos educativos ha experimentado un crecimiento exponencial. Kasneci et al. (2023) realizaron una revisión sistemática del impacto de ChatGPT en la educación, identificando tanto oportunidades como desafíos en la implementación de estas tecnologías. Sus hallazgos sugieren que la IA generativa puede funcionar como asistente pedagógico, pero requiere supervisión humana para garantizar la calidad y relevancia del contenido.

Lo et al. (2023) investigaron específicamente el uso de ChatGPT para la creación de materiales educativos, encontrando que los sistemas de IA pueden generar contenido didáctico de calidad variable, dependiendo de la especificidad de las instrucciones y el dominio de conocimiento. Su estudio reveló que la efectividad de los contenidos generados por IA está directamente relacionada con la claridad de los prompts y la experiencia del usuario en el diseño instruccional.

Sullivan et al. (2023) exploraron las percepciones de educadores universitarios sobre herramientas de IA para la creación de contenido, identificando tanto entusiasmo por las posibilidades innovadoras como preocupaciones sobre la autenticidad y calidad pedagógica del material generado.

#### Monólogos y Humor en Contextos Educativos

En cuanto a la aplicación del humor en entornos educativos, Savage et al. (2017) demostraron que el humor apropiado mejora significativamente la retención de información y reduce la ansiedad estudiantil. Su investigación estableció que el humor contextualizado y relevante para el contenido académico genera mayor impacto que el humor genérico.

Henman (2001) desarrolló un marco teórico para la implementación del humor en la enseñanza universitaria, estableciendo criterios para distinguir entre humor productivo y contraproducente en contextos académicos. Sus hallazgos indican que el humor autodepreciable del instructor y las referencias a experiencias compartidas con los estudiantes resultan más efectivos.

Wanzer et al. (2010) investigaron específicamente el uso de monólogos humorísticos en la enseñanza, encontrando que esta metodología no solo mejora la atención y participación estudiantil, sino que también facilita la comprensión de conceptos complejos mediante analogías y metáforas accesibles.

#### Brecha en el Conocimiento Actual

A pesar de los avances en ambas áreas, existe una notable ausencia de investigación que combine específicamente la IA generativa con la creación de monólogos educativos. Los estudios previos han explorado separadamente la efectividad del humor en la educación y las capacidades de la IA para generar contenido didáctico, pero no han investigado sistemáticamente el potencial de diferentes sistemas de IA para crear guiones de monólogos educativos efectivos.

Además, la literatura carece de evaluaciones comparativas entre diferentes sistemas de IA generativa para tareas pedagógicas específicas, particularmente en formatos innovadores como el stand-up educativo. Esta brecha representa una oportunidad significativa para contribuir al conocimiento sobre nuevas metodologías de enseñanza asistidas por tecnología.

#### Justificación del Estudio

La presente investigación aborda esta brecha mediante la exploración comparativa de las capacidades de dos sistemas de IA generativa (Claude y ChatGPT) para diseñar guiones de monólogos educativos en el contexto de la educación superior en

Pedagogía de la Informática. El estudio evalúa la calidad de los guiones producidos según criterios pedagógicos preestablecidos y analiza su efectividad percibida por estudiantes universitarios.

## 1.2 Preguntas de Investigación y Objetivos

Este estudio se guía por las siguientes preguntas de investigación: ¿En qué medida cumplen los guiones de monólogos educativos generados por IA con los criterios de calidad establecidos para este tipo de recursos pedagógicos? ¿Existen diferencias significativas entre los guiones producidos por diferentes sistemas de IA generativa en cuanto a efectividad educativa percibida? ¿Cuál es la percepción de los estudiantes sobre el aprendizaje logrado a través de monólogos educativos generados por IA?

El objetivo principal es evaluar comparativamente la efectividad de guiones de monólogos educativos generados por dos sistemas de IA generativa diferentes, validados por expertos y presentados a estudiantes universitarios, utilizando métricas específicas de aprendizaje y satisfacción. Los objetivos específicos incluyen: (1) desarrollar y validar una rúbrica específica para evaluar guiones de monólogos educativos tipo stand-up, (2) comparar la calidad de los guiones generados por Claude y ChatGPT según criterios pedagógicos establecidos, y (3) analizar las percepciones estudiantiles sobre la efectividad de ambos tipos de guiones para el aprendizaje en áreas específicas de la informática educativa.

## 2. Materiales y Métodos

### 2.1 Diseño de la investigación

Se implementó un diseño cuasi-experimental comparativo con mediciones post-test para evaluar la efectividad de los guiones generados por dos sistemas de IA diferentes. La investigación se desarrolló en una universidad pública ecuatoriana durante el segundo semestre del año académico 2022-2023, con una duración total de cuatro meses.

### 2.2 Participantes

La muestra estuvo constituida por 40 estudiantes (23 mujeres y 17 hombres) de primer semestre de la carrera de Pedagogía de la Informática, seleccionados mediante muestreo intencional no probabilístico, representando la totalidad de la población disponible. La edad promedio fue de 19.4 años (DE = 1.2).

Adicionalmente, participaron 5 estudiantes de semestres superiores pertenecientes al semillero de investigación (3 mujeres, 2 hombres, edad promedio 21.6 años), quienes actuaron como intérpretes de los monólogos tras dar su consentimiento informado. Para la validación de los guiones participaron 3 expertos: un especialista en artes escénicas con 15 años de experiencia, un pedagogo con doctorado en educación y 12 años de experiencia universitaria, y un experto en tecnología educativa con 10 años de experiencia en diseño instruccional.

### 2.3 Procedimiento

La investigación se realizó en base a las siguientes fases que se muestran:

### Fase 1: Desarrollo y validación de la rúbrica de evaluación

Se diseñó una rúbrica específica para evaluar guiones de monólogos educativos tipo stand-up, considerando criterios pedagógicos y de comunicación efectiva. El proceso de validación incluyó:

**Validación de contenido:** Los tres expertos evaluaron independientemente la pertinencia, claridad y exhaustividad de cada criterio utilizando una escala de 4 puntos (1=no pertinente, 4=muy pertinente). Se calculó el Índice de Validez de Contenido (IVC) para cada ítem, obteniendo valores entre 0.85 y 1.00 (media = 0.94), todos superiores al criterio mínimo de 0.80.

**Prueba piloto:** La rúbrica se aplicó a 8 guiones de monólogos educativos creados previamente por estudiantes de cursos superiores. Esta prueba reveló la necesidad de ajustar los descriptores de dos criterios para mayor claridad.

**Confiabilidad inter-evaluador:** Los tres expertos evaluaron independientemente los mismos 5 guiones de prueba. Se calculó el Coeficiente de Correlación Intraclase (CCI = 0.91, IC 95%: 0.84-0.96), indicando alta concordancia entre evaluadores.

**Consistencia interna:** Se calculó el coeficiente alfa de Cronbach ( $\alpha = 0.89$ ), indicando alta consistencia interna de la rúbrica.

### Fase 2: Generación de guiones

- Primero se definieron seis temáticas específicas relacionadas en las áreas de: Innovación, TICs, Pedagogía, Programación, Robótica Educativa e Informática.
- A continuación, se elaboraron prompts detallados en base a requerimientos para cada sistema de IA (Claude y ChatGPT), incluyendo la rúbrica de evaluación como guía para la generación de contenidos, rúbrica de evaluación que fue validada por experto docente de la carrera de Pedagogía de las Artes y Humanidades, incluyendo: Contexto específico del público objetivo, La rúbrica de evaluación como guía, Extensión aproximada (1500 palabras), Criterios de humor apropiado para contexto académico, Objetivos de aprendizaje específicos por temática
- Se solicitó a cada sistema la creación de 6 guiones (uno por temática), resultando en 12 guiones totales para evaluación.

### Fase 3: Validación por expertos

- Los 12 guiones generados fueron codificados y presentados a los expertos de manera ciega (sin identificar el sistema generador). Cada experto evaluó todos los guiones utilizando la rúbrica validada ver tabla 1. Se seleccionaron los dos guiones mejor evaluados de cada sistema (puntuación  $\geq 45/60$ ) para la fase experimental.

### Fase 4: Preparación e interpretación

- Para este experimento los estudiantes del semillero de investigación recibieron los guiones seleccionados sin conocer qué sistema de Inteligencia artificial generativa los había generado.



- Se realizaron cuatro sesiones de ensayo bajo la supervisión del experto en artes escénicas para garantizar una interpretación adecuada para ejecutarla con el público objetivo, enfocándose en: Técnicas de presentación tipo stand-up, manejo del tiempo y pausas, interacción con audiencia, uso apropiado del humor.
- Los monólogos fueron presentados ante los 40 estudiantes de primer semestre de la carrera de Pedagógica de Informática en dos sesiones diferentes, identificados simplemente como “Guion X” (Claude) y “Guion Y” (ChatGPT).

#### Fase 5: Evaluación y recolección de datos

- Posterior a cada presentación, los estudiantes de primer semestre completaron una encuesta de satisfacción y evaluación del aprendizaje recibido por los monologuistas.
  - La encuesta incluyó ítems cerrados (escala 1-10) para evaluar aspectos generales y aprendizaje por áreas temáticas topadas en el monólogo, así como preguntas abiertas y comparativas, La encuesta incluyó:
- \* Ítems cerrados con escala Likert 1-10 para aspectos generales

- \* Ítems específicos para aprendizaje percibido por áreas temáticas
- \* Preguntas abiertas sobre aspectos valorados
- \* Preguntas comparativas directas entre ambos guiones.

#### 2.4 Instrumentos

Rúbrica de evaluación de guiones: Instrumento validado por experto que evalúa seis criterios: Humor y Enganche, Relevancia del Mensaje, Uso del Lenguaje, Claridad y Coherencia, Interacción con el Público, y Reflexión Final. Cada criterio se califica en una escala que va de 1-10 puntos. Como se muestra en la siguiente tabla:

#### Rango de Calificación para la Validación del Guion

- Rango de 40-50 puntos: Guion completamente factible y altamente recomendado para aplicación en el aprendizaje de los estudiantes.
- Rango de 30-39 puntos: Guion aceptable, requiere algunos ajustes menores para mejorar su eficacia y alineación con los objetivos de aprendizaje.

**Tabla 1.** Rúbrica de Evaluación de Guiones de Monólogos Educativos de Tipo Stand Up

Fuente: Los autores.

| Criterio                   | 8-10  | 5-7  | 1-4  |
|----------------------------|---|--|--|
| Humor y Enganche           | El guion utiliza humor de manera efectiva, enganchando al público desde el inicio y manteniendo el interés durante todo el monólogo.  | El guion tiene algunos elementos humorísticos, pero no siempre son efectivos para captar la atención o mantenerla.   | El humor es escaso o poco efectivo, y no logra captar ni mantener el interés del público.                  |
| Relevancia del Mensaje     | El mensaje del monólogo es relevante y significativo, presentando un tema educativo de manera que invita a la reflexión y aprendizaje.  | El mensaje es algo relevante, aunque hay partes que no están claramente conectadas con el tema educativo principal.  | El mensaje no es relevante o se pierde entre otros elementos del monólogo, dificultando el aprendizaje.    |
| Uso del Lenguaje           | El lenguaje utilizado es adecuado, comprensible y conecta con el público objetivo, favoreciendo la cercanía y la claridad del mensaje.  | El lenguaje es comprensible, aunque en algunas partes resulta complicado o distante del público objetivo.  | El lenguaje es inadecuado o demasiado complejo, lo que dificulta la comprensión y conexión con el público. |
| Claridad y Coherencia      | La estructura del guion es clara, lógica y fácil de seguir, permitiendo que el público comprenda completamente el mensaje.  | La estructura es entendible, aunque presenta algunas partes confusas que dificultan la comprensión.  | La estructura es desordenada y confusa, lo que impide una buena comprensión del mensaje.                   |
| Interacción con el Público | El guion incluye momentos de interacción que fomentan la participación del público y refuerzan el aprendizaje de manera efectiva.   | El guion tiene algunos momentos de interacción, pero no son completamente efectivos para involucrar al público.  | No hay interacción con el público o es muy limitada, reduciendo el impacto del monólogo en el aprendizaje. |
| Reflexión Final            | El guion incluye una reflexión final bien estructurada que ayuda a los estudiantes a consolidar lo aprendido de manera significativa, en base a los ejes de: Innovación, Tics, Pedagogía, Programación, Robótica Educativa e Informática. | La reflexión final está presente, pero no está completamente desarrollada o no es efectiva para consolidar el aprendizaje en base a los ejes de: Innovación, Tics, Pedagogía, Programación, Robótica Educativa e Informática | No hay una reflexión final o es demasiado superficial como para tener un impacto en el aprendizaje         |



- Rango de 20-29 puntos: Guion requiere ajustes importantes, se recomienda revisión completa antes de ser aplicado.

- Rango de 0-19 puntos: Guion no recomendado, no cumple con los estándares básicos para ser utilizado como herramienta educativa.

**Encuesta de satisfacción y aprendizaje:** Instrumento diseñado para esta investigación que evalúa la percepción de los estudiantes sobre diversos aspectos de los monólogos y el aprendizaje adquirido en las seis áreas temáticas, utilizando una escala de 1-10 y preguntas abiertas.

Instrumento estructurado que incluye:

- Datos demográficos: edad, género, experiencia previa con IA
- Evaluación de aspectos generales (5 ítems, escala 1-10): entretenimiento, claridad, conexión, originalidad, calidad
- Aprendizaje percibido por áreas (6 ítems, escala 1-10): una por cada área temática
- Preguntas comparativas: preferencias directas entre guiones
- Preguntas abiertas: aspectos más valorados y conceptos aprendidos

La encuesta fue piloteada con 10 estudiantes de semestres superiores para verificar claridad y tiempo de respuesta (15-20 minutos).

## 2.5 Análisis de datos

Para esto se realizaron análisis descriptivos (medias, desviaciones estándar, frecuencias y porcentajes) e inferenciales (prueba t para muestras relacionadas, tamaño del efecto d de Cohen) para comparar las evaluaciones de los guiones generados por ambos sistemas de IA. Las respuestas a las preguntas abiertas fueron analizadas mediante análisis de contenido temático para identificar patrones que puedan ser recurrentes.

Las técnicas estadísticas incluyeron:

- Prueba t para muestras relacionadas: comparación de puntuaciones medias entre guiones
- Cálculo del tamaño del efecto (d de Cohen): magnitud práctica de las diferencias
- Intervalos de confianza al 95%: precisión de las estimaciones de diferencias
- Prueba de chi-cuadrado: análisis de preferencias categóricas
- Análisis de contenido temático: procesamiento de respuestas abiertas mediante codificación inductiva

El análisis estadístico se realizó con SPSS v.28, estableciendo un nivel de significancia  $\alpha = 0.05$ . Para las respuestas abiertas, dos investigadores realizaron codificación independiente, resolviendo discrepancias mediante discusión y consenso.

## 3. Resultados y Discusión

### 3.1 Validación de expertos

La evaluación de los guiones por parte de los expertos mostró que ambos sistemas de Inteligencia artificial generativa fueron capaces de generar guiones considerados aceptables según la rúbrica establecida. En la tabla 2 presenta las puntuaciones medias y desviaciones estándar otorgadas por los expertos a los guiones seleccionados para su implementación.

**Nota aclaratoria:** Aunque la rúbrica establece rangos de calidad donde “40-50 puntos” corresponde a un “Guion completamente factible y altamente recomendado”, los expertos evaluadores determinaron que puntuaciones superiores a 50 puntos indican un nivel excepcional de calidad, manteniendo la categoría

**Tabla 2.** Rúbrica de Evaluación de Guiones de Monólogos Educativos de Tipo Stand Up  
Fuente: Los autores.

| Criterio evaluado          | Guion X (Claude) | Guion Y (ChatGPT) | Diferencia | ICP 95%    |
|----------------------------|------------------|-------------------|------------|------------|
| Humor y Enganche           | 8.7 (DE= 0,58)   | 7.3 (DE= 0.64)    | +1.4       | [0.8, 2.0] |
| Relevancia del Mensaje     | 9.0 (DE= 0,50)   | 8.3 (DE= 0.57)    | +0.7       | [0.3, 1.1] |
| Uso del Lenguaje           | 8.7 (DE= 0,55)   | 7.7 (DE= 0.61)    | +1.0       | [0.5, 1.5] |
| Claridad y Coherencia      | 9.3 (DE= 0,29)   | 8.0 (DE= 0.53)    | +1.3       | [0.9, 1.7] |
| Interacción con el Público | 8.3 (DE= 0,65)   | 6.7 (DE= 0.76)    | +1.6       | [0.9, 2.3] |
| Reflexión Final            | 9.0 (DE= 0,41)   | 8.0 (DE= 0.48)    | +1.0       | [0.6, 1.4] |
| Puntuación total           | 53.0/60          | 46.0/60           | +7.0       | [4.2, 9.8] |

de “altamente recomendado” pero con mérito adicional. Esta extensión del rango superior permite una mejor discriminación entre guiones de alta calidad, sin modificar los umbrales inferiores de aceptabilidad.

Los resultados de la validación indican que el guion generado por Claude obtuvo una valoración superior en todos los criterios evaluados, con diferencias estadísticamente significativas ( $p < 0.01$ ) en cada dimensión. Las diferencias más pronunciadas se observaron en “Interacción con el Público” (+1.6 puntos, IC 95%: [0.9, 2.3]) y “Humor y Enganche” (+1.4 puntos, IC 95%: [0.8, 2.0]), sugiriendo ventajas específicas en elementos comunicativos y de engagement.

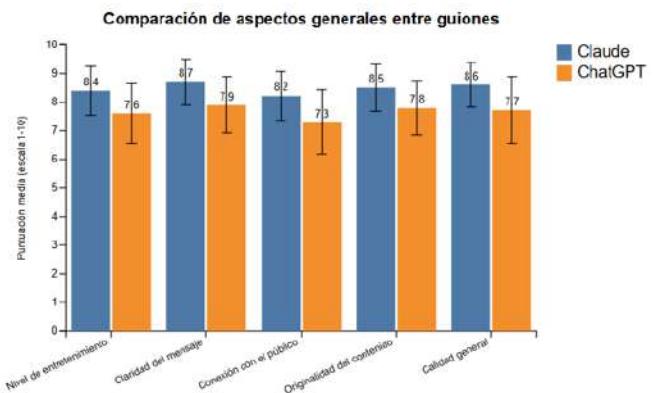


### 3.2 Evaluación de los estudiantes: Aspectos generales

Los datos de la evaluación por parte de los 40 estudiantes revelaron una tendencia consistente favorable hacia el Guion X (Claude) en todas las dimensiones evaluadas, como se presenta en la Tabla 3.

El análisis estadístico revela diferencias significativas entre ambos guiones en todas las dimensiones evaluadas. Los tamaños del efecto observados ( $d = 0.74-0.83$ ) indican diferencias de magnitud grande según los criterios de Cohen, sugiriendo no solo significancia estadística sino también relevancia práctica y pedagógica.

La Figura 1 presenta los resultados de la evaluación por parte de los 40 estudiantes, mostrando una clara tendencia favorable



**Figura 1.** Comparación de las puntuaciones medias y desviaciones estándar en aspectos generales.

Fuente: Los autores.

**Tabla 3.** Puntuaciones medias, desviaciones estándar e intervalos de confianza para aspectos generales  
Fuente: Los autores.

| Aspecto                    | Guion X (Claude) | Guion X (ChatGPT) | Diferencia | ICP 95%      | D de Cohen | p      |
|----------------------------|------------------|-------------------|------------|--------------|------------|--------|
| Nivel de entretenimiento   | 8.4 (DE = 0.85)  | 7.6 (DE = 1.05)   | +0.8       | [0.4, 1.2]   | 0.77       | <0.001 |
| Claridad del mensaje       | 8.7 (DE = 0.79)  | 7.9 (DE = 0.98)   | +0.8       | [0.4, 1.2]   | 0.83       | <0.001 |
| Conexión con el público    | 8.2 (DE = 0.86)  | 7.3 (DE = 1.12)   | +0.9       | [0.5, 1.3]   | 0.83       | <0.001 |
| Originalidad del contenido | 8.5 (DE = 0.82)  | 7.8 (DE = 0.95)   | +0.7       | [0.3, 1.1]   | 0.74       | <0.001 |
| Calidad general            | 8.6 (DE = 0.79)  | 7.7 (DE = 1.18)   | +0.9       | [0.5, 1.3]   | 0.79       | <0.001 |
| Promedio general           | 8.48 (DE = 0.68) | 7.66 (DE = 0.89)  | +0.82      | [0.52, 1.12] | 0.79       | <0.001 |

hacia el Guion X (Claude) en todos los aspectos evaluados.

El análisis estadístico revela diferencias significativas entre ambos guiones. El Guion X (Claude) recibió puntuaciones superiores en todas las dimensiones evaluadas, con diferencias particularmente pronunciadas en “Conexión con el público” ( $M = 8.2$ ,  $DE = 0.86$  vs  $M = 7.3$ ,  $DE = 1.12$ ;  $t(39) = 4.87$ ,  $p < 0.001$ ,  $d = 0.83$ ) y “Calidad general” ( $M = 8.6$ ,  $DE = 0.79$  vs  $M = 7.7$ ,  $DE = 1.18$ ;  $t(39) = 4.53$ ,  $p < 0.001$ ,  $d = 0.79$ ). Estos resultados indican un tamaño del efecto grande, según los criterios de Cohen, lo que sugiere diferencias no solo estadísticamente significativas sino también pedagógicamente relevantes.

### 3.3 Evaluación de los estudiantes: Aprendizaje por áreas temáticas

Respecto al aprendizaje percibido en las seis áreas temáticas, también se observó una ventaja consistente del Guion X (Claude), como se muestra en la Tabla 4.

El análisis estadístico de estos datos revela hallazgos significativos. La diferencia más notable se observa en el área de “Pedagogía”, donde el Guion X (Claude) supera significativamente al Guion Y (ChatGPT) con una diferencia de 1.3 puntos y un tamaño del efecto grande ( $d = 0.91$ ). Esta diferencia es especialmente relevante dado que los participantes

**Tabla 4.** Puntuaciones medias, desviaciones estándar y análisis comparativo del aprendizaje percibido por áreas temáticas.  
Fuente: Los autores.

| Área Temática             | Guion X (Claude)       | Guion Y (ChatGPT)      | Diferencia   | d de Cohen  | p                |
|---------------------------|------------------------|------------------------|--------------|-------------|------------------|
| Innovación                | 8.8 (DE= 0.92)         | 8.1 (DE= 1.14)         | +0.7         | 0.62        | <0.01            |
| TICs                      | 8.7 (DE= 0.83)         | 8.3 (DE= 0.95)         | +0.4         | 0.41        | <0.05            |
| Pedagogía                 | 8.9 (DE= 0.76)         | 7.6 (DE= 1.42)         | +1.3         | 0.91        | <0.001           |
| Programación              | 8.2 (DE= 0.88)         | 7.9 (DE= 0.97)         | +0.3         | 0.35        | <0.05            |
| Robótica Educativa        | 8.3 (DE= 0.94)         | 7.8 (DE= 1.05)         | +0.5         | 0.48        | <0.01            |
| Informática General       | 8.5 (DE= 0.87)         | 8.0 (DE= 1.03)         | +0.5         | 0.49        | <0.01            |
| <b>Promedio por Áreas</b> | <b>8.57 (DE= 0.72)</b> | <b>7.95 (DE= 0.58)</b> | <b>+0.62</b> | <b>0.56</b> | <b>&lt;0.001</b> |

son estudiantes de Pedagogía de la Informática, lo que sugiere que el guion generado por Claude logró conectar de manera más efectiva con los conceptos fundamentales de su disciplina. Las áreas de “Innovación” ( $d = 0.62$ ) y “TICs” ( $d = 0.41$ ) también muestran diferencias considerables a favor del Guion X, con tamaños del efecto moderados, lo que refuerza la tendencia general observada.

### 3.4 Preferencias comparativas

La evaluación directa de preferencias por parte de los estudiantes refuerza los resultados cuantitativos anteriores. El 70% ( $n = 28$ ) indicó que el Guion X (Claude) fue más efectivo para su aprendizaje, frente al 20% ( $n = 8$ ) que prefirió el Guion Y (ChatGPT) y un 10% ( $n = 4$ ) que los consideró igualmente efectivos ( $\chi^2 = 27.2$ ,  $p < 0.001$ ). Este patrón se repite en las dimensiones de entretenimiento (62.5% vs 25%,  $\chi^2 = 18.4$ ,  $p < 0.001$ ) y claridad (65% vs 22.5%,  $\chi^2 = 21.6$ ,  $p < 0.001$ ), siempre favoreciendo al Guion X con diferencias estadísticamente significativas.

### 3.5 Análisis de respuestas abiertas

Para el **Guion X (Claude)**, los estudiantes valoraron especialmente cuatro aspectos clave que ejemplifican su enfoque pedagógico distintivo:

1. Contextualización del humor: Los elementos humorísticos surgían orgánicamente de situaciones reconocibles para los estudiantes. Como expresó un participante:

“El humor no parecía forzado, sino que surgía naturalmente de situaciones relacionadas con la informática que todos hemos vivido.” Este enfoque representa lo que Garner (2020) denomina “humor pedagógicamente relevante”, que establece conexiones significativas con el contenido académico.

2. Ejemplos situados en la experiencia estudiantil: Claude demostró capacidad para relacionar conceptos abstractos con experiencias cotidianas. Extracto representativo del guion:

“Imaginen que están programando como quien prepara una receta. No se trata solo de echar ingredientes al azar; es entender que cada línea de código, como cada ingrediente, cumple una función específica y se relaciona con el todo. Cuando comprendemos esta relación, dejamos de memorizar comandos para empezar a ‘pensar en código’.”

3. Estrategias de interacción auténticas Las preguntas propuestas fueron percibidas como genuinas. Ejemplo del guion:

“¿Cuántos de ustedes han sentido esa frustración cuando no encuentran el error en su código? ¿Ven ese momento como un fracaso o como parte del proceso de aprendizaje? [pausa para respuestas] Interesante... la robótica educativa nos enseña precisamente que el error es información valiosa, no un callejón sin salida.”

4. Estructura metacognitiva: Promovía reflexión sobre el propio aprendizaje:

“Lo fascinante de la pedagogía informática no es solo enseñar a usar herramientas, sino desarrollar una forma de pensar que conecta la lógica computacional con la creatividad humana. Lo que hemos explorado hoy no son solo conceptos técnicos, sino formas de construir puentes entre el pensamiento algorítmico y las necesidades educativas reales.”

Por su parte, el **Guion Y (ChatGPT)** mostró fortalezas diferentes, siendo más valorado por aspectos como:

1.Precisión informativa: “Proporcionaba datos interesantes sobre la historia de la informática educativa, fechas y referencias que complementaban la explicación.”

2. Estructura didáctica convencional: “La organización temática facilitaba seguir el hilo del monólogo, con una introducción clara, desarrollo y conclusión.”

3.Manejo adecuado de terminología: “Usaba términos técnicos pero los explicaba bien, definiendo conceptos complejos antes de aplicarlos.”



Estos hallazgos sugieren que mientras ChatGPT tiende a priorizar la estructura didáctica convencional y la precisión informativa, Claude logra una mayor integración de elementos pedagógicos avanzados como el aprendizaje situado, la metacognición y la construcción de relevancia personal para el estudiante.

### 3.6 Discusión

Los resultados de la investigación revelan patrones significativos sobre las capacidades de dos sistemas de Inteligencia artificial generativa para crear contenido educativo efectivo en formato de monólogo stand-up. Más allá de las diferencias cuantitativas observadas, es fundamental profundizar en los factores subyacentes que explican la mayor efectividad pedagógica del sistema Claude con respecto a ChatGPT.

#### 3.6.1 Arquitectura narrativa y principios pedagógicos

El análisis detallado de los guiones sugiere que Claude integra de manera más efectiva principios pedagógicos contemporáneos en su estructura narrativa. Tres elementos fundamentales se destacan:

**1. Aprendizaje situado:** Claude construye sistemáticamente puentes entre conceptos abstractos y experiencias concretas. Por ejemplo, al explicar algoritmos, utiliza metáforas como “la receta de cocina familiar que se va perfeccionando con cada generación”, estableciendo conexión con experiencias previas de los estudiantes. Este enfoque refleja lo que Lave y Wenger (1991) denominan “cognición situada”, donde el aprendizaje se contextualiza en prácticas auténticas.

**2. Andamiaje cognitivo:** El guion de Claude implementa una progresión cuidadosa que construye comprensión paso a paso. Comienza con analogías familiares, avanza hacia definiciones formales, y culmina con aplicaciones complejas.

Este enfoque sigue el principio de “andamiaje” propuesto por Bruner (1978), proporcionando apoyos temporales que se retiran gradualmente a medida que aumenta la competencia del estudiante.

**3. Interacción dialógica:** A diferencia del enfoque más expositivo de ChatGPT, Claude incorpora elementos de diálogo socrático, planteando preguntas que invitan a la reflexión crítica.

Estas preguntas no son meramente retóricas, sino que crean espacios cognitivos para que los estudiantes construyan conexiones personales con el material.

Estos elementos no aparecen como componentes aislados, sino integrados en una estructura narrativa coherente que mantiene el humor como vehículo para el aprendizaje, no como mero adorno.

#### 3.6.2 Factores tecnológicos explicativos

Las diferencias observadas pueden atribuirse a variaciones en las arquitecturas de entrenamiento y procesamiento de instrucciones de ambos sistemas. Claude demostró mayor capacidad para mantener coherencia narrativa y contextualización educativa, mientras que ChatGPT mostró fortalezas en estructura didáctica convencional.

Como señalan Kasneci et al. (2023), los sistemas de IA generativa muestran fortalezas variables según el dominio y tipo de tarea, lo que sugiere que no existe una superioridad generalizable de un sistema sobre otro, sino aptitudes diferenciadas para aplicaciones específicas..

#### 3.6.3 Implicaciones pedagógicas

Los hallazgos de este estudio tienen importantes implicaciones para la práctica educativa:

**1. Complementariedad con la labor docente:** Los monólogos generados por IA pueden servir como punto de partida para que los educadores desarrollen materiales didácticos innovadores. Sin embargo, el valor añadido del criterio pedagógico humano sigue siendo insustituible para adaptar estos contenidos a contextos específicos.

**2. Personalización del aprendizaje:** La capacidad de generar rápidamente contenido educativo en formatos atractivos abre posibilidades para personalizar materiales según las necesidades e intereses específicos de diferentes grupos de estudiantes.

**3. Nuevos formatos pedagógicos:** El éxito del formato monólogo tipo stand-up sugiere que existen oportunidades para explorar formatos educativos no convencionales que combinan entretenimiento y rigor académico.

#### 3.6.4 Limitaciones del estudio

Es importante reconocer varias limitaciones que contextualizan los resultados:

**1. Tamaño y homogeneidad de la muestra:** Con 40 participantes, todos de la misma carrera y universidad, los resultados pueden no ser generalizables a poblaciones más amplias o diversas.

**2. Sesgo de familiaridad:** Los estudiantes pueden tener diferentes niveles de exposición previa a sistemas de IA, lo que podría influir en sus percepciones.

**3. Subjetividad en la percepción del humor:** La efectividad del humor es altamente contextual y culturalmente dependiente, lo que limita la generalización de estos hallazgos a otros contextos culturales.

**4. Variables no controladas:** Factores como la habilidad interpretativa de los estudiantes que presentaron los monólogos o las condiciones específicas de presentación pueden haber influido en los resultados.

**5. Temporalidad de la tecnología:** Dado el rápido avance de los sistemas de IA, es probable que las capacidades de ambos sistemas evolucionen significativamente en corto tiempo, lo que podría alterar los resultados en estudios futuros.

Estas limitaciones, lejos de invalidar los hallazgos, ofrecen direcciones valiosas para investigaciones futuras que podrían abordar aspectos como el impacto a largo plazo en el aprendizaje, la evaluación objetiva de conocimientos adquiridos, o la comparación con contenido generado enteramente por humanos.

#### 4. Conclusiones

Este trabajo de investigación demuestra el potencial de la Inteligencia artificial generativa como una herramienta para crear guiones para monólogos educativos a nivel de educación universitaria.

Los hallazgos sugieren que tanto las Inteligencia artificial generativas Claude como ChatGPT fueron capaces de producir guiones considerados viables y recomendables por todos los criterios establecidos, aunque había diferencias notables en la efectividad percibida de cada sistema.

Los expertos y estudiantes atribuyeron consistentemente mejores calificaciones a los guiones generados por Claude, especialmente en las áreas de conexión emocional, claridad del mensaje y posibilidad de generar reflexión profunda en disciplinas pedagógicas. Estas diferencias implican que, al elegir herramientas de IA para la generación de contenido educativo, los educadores pueden tener que tener en cuenta las fortalezas particulares de cada sistema en relación con sus objetivos de enseñanza.

La respuesta positiva de los estudiantes hacia los monólogos educativos generados (el 90% recomendaría su uso en otros cursos) demuestra la efectividad de este enfoque metodológico como una adición a las técnicas de enseñanza tradicionales.

Un estilo de monólogo de comedia, cuando está bien elaborado, tiene el potencial de combinar efectivamente el humor con el aprendizaje serio, mejorando la comprensión y retención de conceptos difíciles por parte de los estudiantes.

En base a los resultados de las encuestas se pudo decir que en términos de aspectos generales, el Guion X recibió puntuaciones medias superiores en todos los criterios evaluados.

La diferencia más significativa se pudo observar en “Conexión con el público” ( $M = 8.2$  vs  $M = 7.3$ ,  $d = 0.83$ ) y “Calidad general” ( $M = 8.6$  vs  $M = 7.7$ ,  $d = 0.79$ ), ambas con un tamaño del efecto considerado grande ( $d > 0.8$ ).

Dentro de la evaluación que se realizó en el ‘Aprendizaje por áreas temáticas’, la diferencia más notable se pudo observar

en el área de “Pedagogía”, donde el Guion X (Claude) superó significativamente al Guion Y (ChatGPT) con una diferencia marcada de 1.3 puntos y un tamaño del efecto grande ( $d = 0.91$ ). Las áreas de “Innovación” y “TICs” también mostraron diferencias considerables a favor del Guion X, aunque con tamaños del efecto moderados.

Los resultados del estudio pueden ser útiles para los profesores, especialistas en diseño instruccional y desarrolladores de sistemas de Inteligencia artificial. Para los docentes, proporciona una herramienta metodológica innovadora que puede enriquecer sus prácticas de enseñanza, notablemente en aquellos casos donde el humor contextualizado puede ayudar a comprender conceptos abstractos. Para los diseñadores, resalta áreas particulares que los sistemas existentes descuidan.

Esta investigación propone a futuro el uso nuevas herramientas y tecnologías de IA como sistemas emergentes para Evaluar las capacidades de sistemas de IA más recientes (GPT-4, Bard, LLaMA 2, PaLM) y herramientas especializadas en educación conforme estén disponibles, la IA multimodal: Investigar sistemas que combinan generación de texto con imágenes, audio y video para crear contenido educativo más rico y diverso, la personalización avanzada: Explorar sistemas de IA que puedan adaptar automáticamente el estilo, humor y contenido según perfiles individuales de estudiantes o características grupales específicas.

Este estudio representa un primer paso en la exploración sistemática del potencial de la IA generativa para revolucionar la creación de contenido educativo innovador. Los resultados sugieren que estamos en el umbral de una nueva era donde la tecnología puede amplificar significativamente la creatividad pedagógica, permitiendo a los educadores acceder a herramientas poderosas para crear experiencias de aprendizaje más atractivas y efectivas. Sin embargo, los hallazgos también subrayan que la tecnología por sí sola no garantiza la calidad educativa. La supervisión pedagógica experta, la adaptación contextual y la evaluación continua siguen siendo elementos insustituibles para asegurar que las innovaciones tecnológicas se traduzcan en mejoras reales del aprendizaje estudiantil.

Los hallazgos de este estudio abren múltiples líneas de investigación prometedoras que ampliarían significativamente el conocimiento en esta área emergente. Se recomienda desarrollar estudios longitudinales que evalúen el impacto a largo plazo de los monólogos educativos generados por IA en la retención de conocimientos y el rendimiento académico durante semestres completos.

La investigación se beneficiaría de muestras más diversas que incluyan diferentes carreras, niveles educativos y contextos culturales para evaluar la generalización de estos hallazgos más allá de la Pedagogía de la Informática. Asimismo, resulta esencial explorar nuevos géneros discursivos como debates estructurados, simulaciones históricas interactivas, storytelling educativo y dramatizaciones científicas para determinar la



versatilidad de la IA generativa en formatos pedagógicos diversos. Finalmente, la rápida evolución tecnológica demanda investigaciones con herramientas emergentes de IA como GPT-4, sistemas multimodales que integren texto-imagen-audio, y plataformas especializadas en educación, permitiendo comparaciones actualizadas de capacidades y la identificación de nuevas oportunidades para la innovación pedagógica asistida por inteligencia artificial.

#### Agradecimientos

Agradecimiento a estudiantes y docentes de la Carrera de Pedagogía de las Ciencias Experimentales de la Informática en Facultad de Educación de la Universidad Nacional de Chimborazo, así como a estudiantes del semillero de investigación de la Carrera.

#### Contribución de los autores

**Christiam Xavier Núñez Zavala:** Conceptualización, Investigación, Software. **Jorge Noe Silva Castillo:** Metodología, Recursos, Análisis formal, Redacción – Diseño editorial. **Cristhy Nataly Jimenez Granizo:** Metodología, Validación, Redacción – Revisión y Edición, Supervisión. **Hernán Ramiro Pailiacho Yucta:** Conceptualización, Redacción – Borrador Original, Financiación

#### Conflictos de interés

Los autores no tienen conflictos de intereses.

#### Referencias bibliográficas

- Alcalá, H. C. (2008). Los monólogos. Algunas notas para su análisis. *Oralia: Análisis del Discurso Oral*, 11, 421–436.
- Anthropic. (2023). Claude 2.1: Introducing 200K context windows, reduced rates of model hallucination, and system prompts. *Anthropic*. <https://www.anthropic.com/news/clause-2-1>
- Bai, Y., Jones, A., Ndousse, K., Askell, A., Chen, A., DasSarma, N., Drain, D., Fort, S., Ganguli, D., Henighan, T., Joseph, N., Kadavath, S., Kernion, J., Conerly, T., ElShowk, S., Elhage, N., Hatfield-Dodds, Z., Hernandez, D., Hume, T., ... Kaplan, J. (2022). Training a helpful and harmless assistant with reinforcement learning from human feedback (arXiv:2204.05862). *arXiv*. <https://doi.org/10.48550/arXiv.2204.05862>
- Basulto-González, G., & Jorge-Hechavarria, R. (2018). El enfoque sociocultural-profesional en la formación de profesores. *Luz*, 17(4), 57–66.
- Berk, R. A. (2007). Humor as an instructional defibrillator: From triage to technique. *The Journal of Health Administration Education*, 31(1), 1–10.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., & Amodei, D. (2022). Finetuned language models are zero-shot learners (arXiv:2109.01652). *arXiv*. <https://arxiv.org/abs/2109.01652>
- Bruner, J. (1978). The role of dialogue in language acquisition. In A. Sinclair, R. J. Jarvelle, & W. J. M. Levelt (Eds.), *The child's conception of language* (pp. 241–256). Springer-Verlag.
- Comparato, D. (2014). De la creación al guión: Arte y técnica de escribir para cine y televisión. *Simplíssimo*.
- Escrivá, V. M., García, P. S., & Navarro, M. D. F. (2002). Procesos cognitivos y emocionales predictores de la conducta prosocial y agresiva: La empatía como factor modulador. *Psicothema*, 14(2), 227–232.
- Garner, R. L. (2020). *Humor in pedagogy: How ha-ha can lead to aha!* *College Teaching*, 68(1), 48–57. <https://doi.org/10.1080/87567555.2019.1694072>
- González-Encinas, J., Saúl, L. Á., & García-Martínez, J. (2019). Revisión de la autocaracterización: Una técnica narrativa constructivista. *Acción Psicológica*, 16(1), 105–128. <https://doi.org/10.5944/ap.16.1.22192>
- Guilera, L. (2020). *Anatomía de la creatividad*. Marge Books.
- Henman, L. D. (2001). Humor as a coping mechanism: Lessons from POWs. *Humor*, 14(1), 83–94. <https://doi.org/10.1515/humr.14.1.83>
- Hernández Rosado, M., Lluesma Rojas, M. de la C., & De Veras Olivera, B. (2019). Hacia una comunicación eficaz. *Revista Cubana de Educación Superior*, 38(2). [http://scielo.sld.cu/scielo.php?script=sci\\_abstract&pid=S0257-43142019000200006](http://scielo.sld.cu/scielo.php?script=sci_abstract&pid=S0257-43142019000200006)
- Hernández-Calderón, K. A., & Lesmes-Silva, A. K. (2018). La escucha activa como elemento necesario para el diálogo. *Revista Convicciones*, 5(9), 83–87.
- Izquierdo, R. M. (2024). Evaluación de ChatGPT como asistente en el proceso de creación de un guion de cortometraje de ficción. *Revista de la Asociación Española de Investigación de la Comunicación*, 11(Especial), e07.

- <https://doi.org/10.24137/raec.11.e07>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kumar, V., Boulanger, D., Seanosky, J., Panés-Ruiz, K., Géraud, M., Lafontaine, W., & Kinshuk. (2023). Leveraging emotion detection in educational chatbots: A systematic review. *Journal of Educational Computing Research*, 61(2), 439–472. <https://doi.org/10.1177/07356331221123456>
- Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge University Press.
- Lo, C. K., Hew, K. F., & Chen, C. (2023). Toward a set of design principles for mathematics flipped classrooms: A synthesis of research in mathematics education. *Educational Research Review*, 39, 100504. <https://doi.org/10.1016/j.edurev.2023.100504>
- Morales-Chan, M. A. (2023). Explorando el potencial de ChatGPT: Una clasificación de prompts efectivos para la enseñanza. <https://biblioteca.galileo.edu/xmlui/handle/123456789/1348>
- Reyna Treviño, C. A. (2020). La autorregulación emocional y cognitiva a favor del aprendizaje. *Educando para Educar*, 38(septiembre 2019–febrero 2020), 109–120.
- Savage, B. M., Lujan, H. L., Thipparthi, R. R., & DiCarlo, S. E. (2017). Humor, laughter, learning, and health! A brief review. *Advances in Physiology Education*, 41(3), 341–347. <https://doi.org/10.1152/advan.00030.2017>
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*, 6(1), 31–40. <https://doi.org/10.37074/jalt.2023.6.1.17>
- Uriol, D. D. S., & Cueva, J. J. C. (2022). Estrategias metodológicas para promover el pensamiento crítico en los estudiantes. *Ciencia Latina Revista Científica Multidisciplinaria*, 6(3), Article 3. [https://doi.org/10.37811/cl\\_rcm.v6i3.2434](https://doi.org/10.37811/cl_rcm.v6i3.2434)
- Veletsianos, G. (2012). How do learners respond to pedagogical agents that deliver social-oriented non-task messages? Impact on student learning, perceptions, and experiences. *Computers in Human Behavior*, 28(3), 749–756. <https://www.sciencedirect.com/science/article/pii/S0747563211002032>
- Wanzer, M. B., Frymier, A. B., Wojtaszczyk, A. M., & Smith, T. (2006). Appropriate and inappropriate uses of humor by teachers. *Communication Education*, 55(2), 178–196. <https://doi.org/10.1080/03634520600566132>
- Motivation of High School Students. *Electronics*, 12(22), 4715. <https://doi.org/10.3390/electronics12224715>
- Cabascango Trávez, G. (2023). El uso de la realidad aumentada en la enseñanza de ciencias: Un enfoque integrador en educación secundaria. *Revista Científica Kosmos*, 2(1), 39–50. <https://doi.org/10.62943/RCK.V2N1.2023.43>
- Cabero Almenara, J., Vázquez Cano, E., Villota Oyarvide, W. R., & López Meneses, E. (2021). La innovación en el aula universitaria a través de la realidad aumentada. *Educare*, 25(3), 1–17. <https://doi.org/10.7440/res64.2018.03>
- Cabero-Almenara, J., Vázquez-Cano, E., Villota-Oyarvide, W. R., & López-Meneses, E. (2021). La innovación en el aula universitaria a través de la realidad aumentada. Análisis desde la perspectiva del estudiantado español y latinoamericano. *Revista Electrónica Educare*, 25(3), 1–17. <https://doi.org/10.15359/REE.25-3.1>
- Calderón Imbaquingo, Z. L., Sam Anlas, C. A., Hubel Solis, B., & Huancollo Quispe, D. (2024). Integración de realidad aumentada en la enseñanza de lenguaje en la educación superior: transformando la experiencia de aprendizaje. *Reincisol*, 3(6), 3390–3414. [https://doi.org/10.59282/reincisol.V3\(6\)3390-3414](https://doi.org/10.59282/reincisol.V3(6)3390-3414)
- Cárdenas Castellanos, J., Camacho Vázquez, M. R., Adaile Benítez, N. T., & Pérez Méndez, J. (2024). Realidad aumentada en educación universitaria. *Emerging Trends in Education*, 7(13), 117–128. <https://doi.org/10.19136/ETIE.A7N13.6300>
- Carrión, J., & Serrano, V. (2021). Revisión sistemática de literatura: características y funcionamiento respecto a los modelos BERT y SQuAD. *CEDAMAZ: Revista del Centro de Estudio y Desarrollo de la Amazonía*, 11(1), 79–86.
- Del Zulia, U., Luis, V., & Arredondo, A. L. (2020). Realidad Aumentada Móvil: Una estrategia pedagógica en el ámbito universitario. *Revista Técnica de Ingeniería de la Universidad del Zulia*, 43(3), 142–149. <https://doi.org/10.22209/rt.v43n3a04>
- Dengel, A., Iqbal, M. Z., Gafe, S., & Mangina, E. (2022). A review on augmented reality authoring toolkits for education. *Frontiers in Virtual Reality*, 3. <https://doi.org/10.3389/FRVIR.2022.798032>
- Geerlings-Batt, J., Tillett, C., Gupta, A., & Sun, Z. (2022). Enhanced visualisation of normal anatomy with potential use of augmented reality superimposed on three-dimensional printed models. *Micromachines*, 13(10), 1701. <https://doi.org/10.3390/MI13101701>



Gonzalez, A. A., Lizana, P. A., Pino, S., Miller, B. G., & Merino, C. (2020). Augmented reality-based learning for the comprehension of cardiac physiology in undergraduate biomedical students. *Advances in Physiology Education*, 44(3), 314–322. <https://doi.org/10.1152/ADVAN.00137.2019>

González Pérez, A., & Cerezo Cortijo, I. (2020). Implicaciones pedagógicas de la realidad aumentada para la mejora de la enseñanza de las ciencias en primaria. *Revista Interuniversitaria de Investigación En Tecnología Educativa*, 1–16. <https://doi.org/10.6018/riite.444961>

González Vidal, I., Cebreiro López, B., & Casal Otero, L. (2021). Nuevas competencias digitales en estudiantes potenciadas con el uso de Realidad Aumentada. Estudio Piloto. RIED. *Revista Iberoamericana de Educación a Distancia*, 24(1). <https://doi.org/10.5944/ried.24.1.27501>

Grodotzki, J., Müller, B. T., & Tekkaya, A. E. (2023). Introducing a general-purpose augmented reality platform for the use in engineering education. *Advances in Industrial and Manufacturing Engineering*, 6, 100116. <https://doi.org/10.1016/J.AIME.2023.100116>

Handoyo, K. J., Wisnuwardana, C. J., Austen, A., & Permana, F. (2024). Molecule World: Enhancing chemistry education through web-based augmented reality using Assemblr. *Procedia Computer Science*, 245(C), 1249–1258. <https://doi.org/10.1016/J.PROCS.2024.10.354>

Hernández, D., Bottner, E., Cataldo, F., & Zaragoza, E. (2021). Aplicación de Realidad Aumentada para Laboratorios de Química. *Educación Química*, 32(3), 30–37. <https://doi.org/10.22201/FQ.18708404E.2021.3.68129>

Hurtado-Mazeyra, A., Condori-Yucra, N., Ponce-Alvarez, E., Limaymanta, C. H., & Suárez-Guerrero, C. (2024). Uso didáctico de la Realidad Aumentada en la Educación Infantil: Una revisión sistemática. *Revista Complutense de Educación*, 35(3), 515–528. <https://doi.org/10.5209/rced.85815>

Jang, J., Ko, Y., Shin, W. S., & Han, I. (2021). Augmented Reality and Virtual Reality for Learning: An Examination Using an Extended Technology Acceptance Model. *IEEE Access*, 9, 6798–6809. <https://doi.org/10.1109/ACCESS.2020.3048708>

Kamińska, D., Zwoliński, G., Laska-Leśniewicz, A., Raposo, R., Vairinhos, M., Pereira, E., Urem, F., Ljubić Hinić, M., Haamer, R. E., & Anbarjafari, G. (2023). Augmented Reality: Current and New Trends in

Education. *Electronics (Switzerland)*, 12(16). <https://doi.org/10.3390/electronics12163531>

Kaviyaraj, R., & Uma, M. (2021). A Survey on Future of Augmented Reality with AI in Education. Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021, 47–52. <https://doi.org/10.1109/ICAIS50930.2021.9395838>

Kaviyaraj, R., & Uma, M. (2022). Augmented Reality Application in Classroom: An Immersive Taxonomy. Proceedings - 4th International Conference on Smart Systems and Inventive Technology, ICSSIT 2022, 1221–1226. <https://doi.org/10.1109/ICSSIT53264.2022.9716325>

Lancheros-Bohorquez, W. F., & Vesga-Bravo, G. J. (2024). Uso de la realidad aumentada, la realidad virtual y la inteligencia artificial en educación secundaria: una revisión sistemática. *Revista de Investigación, Desarrollo e Innovación*, 14(1), 95–110. <https://doi.org/10.19053/UPTC.20278306.V14.N1.2024.17537>

Leal Aragón, L. (2020). Producción de recursos didácticos para el aula de matemáticas de Secundaria con realidad aumentada. *Innovación Educativa*, 30, 185–198. <https://doi.org/10.15304/ie.30.6905>

Lin, W., Lo, W. T., & Yueh, H. P. (2022). Effects of learner control design in an AR-based exhibit on visitors' museum learning. *PLOS ONE*, 17(10), e0274826. <https://doi.org/10.1371/journal.pone.0274826>

Lucero Baldevenites, E. V. (2024). Transformando la educación: IA y realidades aumentadas y virtual en la formación docente. *European Public & Social Innovation Review*, 9, 1–16. <https://doi.org/10.31637/epsir-2024-854>

Márquez Díaz, J. E., & Morales Espinosa, L. A. (2020). Realidad aumentada como herramienta de apoyo al aprendizaje de las funciones algebraicas y trascendentes. *Revista Educación En Ingeniería*, 15(29), 34–41. <https://doi.org/10.26507/rei.v15n29.1037>

Martínez, J. O. (2024). La realidad aumentada y realidad virtual en la enseñanza matemática: educación inclusiva y rendimiento académico. Edutec. *Revista Electrónica de Tecnología Educativa*, 88(88), 62–76. <https://doi.org/10.21556/edutec.2024.88.3133>

Martínez, O. M., Mejía, E., Ramírez, W. R., & Rodríguez, T. D. (2021). Incidencia de la realidad aumentada en los procesos de aprendizaje de las funciones matemáticas. *Información Tecnológica*, 32(3), 3–14. <https://doi.org/10.1109/infotech.2021.9400003>

- org/10.4067/S0718-07642021000300003
- Masood, Z., Qabool, H., Fida, M., & Sukhia, R. H. (2024). Exploring the knowledge and awareness on applications of virtual reality and augmented reality technology among dental healthcare professionals – a cross-sectional survey. *Journal of the Pakistan Medical Association*, 74(4), S10–S16. <https://doi.org/10.47391/JPMA.AKU-9S-03>
- Menteclara, F., Roma, A., & Cecilia, M. (2022). La realidad virtual en escuelas hospitalarias. *Revista Científica Arbitrada de La Fundación MenteClara*, 7, 1–24. <https://doi.org/10.32351/rca.v7.290>
- Moreno Fuentes, E., Hidalgo Navarrete, J., Burgos Bolós, C., & Blanca de la Paz, S. de la. (2021). Aprendizaje integrado y colaborativo de ciencias a través de la realidad aumentada en educación infantil. *Etic@net: Revista Científica Electrónica de Educación y Comunicación en la Sociedad del Conocimiento*, 21(1), 214–229. <https://doi.org/10.30827/eticanet.v21i1.17012>
- Muzyleva, I., Yazykova, L., Gorlach, A., & Gorlach, Y. (2021). Augmented and virtual reality technologies in education. Proceedings - 2021 1st International Conference on Technology Enhanced Learning in Higher Education, TELE 2021, 99–103. <https://doi.org/10.1109/TELE52840.2021.9482568>
- Nadeem, M., Lal, M., Cen, J., & Sharsheer, M. (2022). AR4FSM: Mobile Augmented Reality Application in Engineering Education for Finite-State Machine Understanding. *Education Sciences*, 12(8), 555. <https://doi.org/10.3390/educsci12080555>
- Negi, S. K. (2024). Exploring the Impact of Virtual Reality and Augmented Reality Technologies in Sustainability Education on Green Energy and Sustainability Behavioral Change: A Qualitative Analysis. *Procedia Computer Science*, 236, 550–557. <https://doi.org/10.1016/j.procs.2024.05.065>
- Nevrelova, N., Korenova, L., Lavicza, Z., Bruzkova, N., & Schmid, A. (2024). Enhancing digital literacy in primary education through augmented reality. *Frontiers in Education*, 9, 1390491. <https://doi.org/10.3389/feduc.2024.1390491>
- Nunes, M., Adão, T., Shahrabadi, S., Capela, A., Carneiro, D., Branco, P., Magalhães, L., Morais, R., & Peres, E. (2024). ARPocketLab—A Mobile Augmented Reality System for Pedagogic Applications. *Computers*, 13(6), 148. <https://doi.org/10.3390/computers13060148>
- Núñez Zavala, C. X., Isín Vilema, M. D., Jiménez Granizo, C. N., & Peñaflie Barros, G. O. (2023). Inteligencia artificial y realidad aumentada como herramientas innovadoras en el desarrollo de cuentos educativos. *Informática y Sistemas: Revista de Tecnologías de la Informática y las Comunicaciones*, 32, 135–146. <https://doi.org/10.6035/artseduca.6272>
- Comunicaciones, 7(2), 87–95. <https://doi.org/10.33936/isrtic.v7i2.6156>
- Pabuçcu-Aklış, A. (2024). Using innovative technology tools in organic chemistry education: bibliometric analysis. *Chemistry Teacher International*, 7(1), 141–156. <https://doi.org/10.1515/cti-2024-0055>
- Pérez-Muñoz, S., Castaño Calle, R., Morales Campo, P. T., & Rodríguez-Cayetano, A. (2024). A Systematic Review of the Use and Effect of Virtual Reality, Augmented Reality and Mixed Reality in Physical Education. *Information*, 15(9), 582. <https://doi.org/10.3390/info15090582>
- Radu, I., & Schneider, B. (2023). Designing augmented reality for makerspaces: Guidelines, lessons and mitigation strategies from 5+ years of AR educational projects. *Computers & Education: X Reality*, 2, 100026. <https://doi.org/10.1016/j.cexr.2023.100026>
- Real-Fernandez, A., Molina-Carmona, R., & Llorens-Largo, F. (2021). How suitable is for learners an autonomous, interactive and dynamic learning model? Proceedings of 2021 World Engineering Education Forum/Global Engineering Deans Council, WEEF/GEDC 2021, 617–623. <https://doi.org/10.1109/WEEF-GEDC53299.2021.9657378>
- Rebelo, C. M., Deiró, G. F., Knuutila, H. K., Moreira, L. C. de S., & Nogueira, I. B. R. (2024). Augmented reality for chemical engineering education. *Education for Chemical Engineers*, 47, 30–44. <https://doi.org/10.1016/j.ece.2024.04.001>
- Ribič, L., & Devetak, I. (2024). Augmented reality in developing students' understanding of chemistry triplet: A systematic literature review. *Chemistry Teacher International*, 7(1), 157–172. <https://doi.org/10.1515/cti-2024-0060>
- Rivas Rebaque, B., Gértrudix Barrio, F., & Gértrudix-Barrio, M. (2021). Análisis sistemático sobre el uso de la realidad aumentada en Educación Infantil. Edutec. *Revista Electrónica de Tecnología Educativa*, 76, 53–73. <https://doi.org/10.21556/edutec.2021.76.2053>
- Rodrigues, R., & Pombo, L. (2024). The potential of a mobile augmented reality game in education for sustainability: Report and analysis of an activity with the EduCITY App. *Sustainability*, 16(21), 9357. <https://doi.org/10.3390/su16219357>
- Rodríguez López, M. (2022). Desarrollo del aprendizaje basado en proyectos con realidad aumentada en educación secundaria para mejorar rendimientos en el aula de música. *Artseduca*, 32, 135–146. <https://doi.org/10.6035/artseduca.6272>
- Sacoto, F., Guillen, V. M., & Hevia Artíme, I. (2024). El aprendizaje del lenguaje, un motor de desarrollo: Revisión sistemática de literatura. *Praxis Pedagógica*,



- 24(36), 102–131. <https://doi.org/10.26620/uniminuto.praxis.24.36.2024.102-131>
- Salgado Reveles, M. A. (2023). Los efectos de la realidad virtual y la realidad aumentada en las actitudes hacia la ciencia en alumnos mexicanos de nivel primaria. *PAAKAT: Revista de Tecnología y Sociedad*, 13(25), 1–30. <https://doi.org/10.32870/pk.a13n25.804>
- Sánchez Hayman, L. H., Morales Tamayo, Y., & Trujillo Ronquillo, D. F. (2024). Revisión sistemática de literatura sobre la incidencia de la tecnología CNC en la Industria 4.0. *Revista Científica y Tecnológica UPSE*, 11(2), 145–155. <https://doi.org/10.26423/rectu.v11i2.801>
- Singh, S., Kaur, A., & Gulzar, Y. (2024). The impact of augmented reality on education: A bibliometric exploration. *Frontiers in Education*, 9, 1458695. <https://doi.org/10.3389/feduc.2024.1458695>
- Tebes, G., Becker, P., Peppino, D., & Olsina, L. (2019, April). Especificación del modelo de proceso para una revisión sistemática de literatura. Proceedings of Curran Associates Publisher 2019. <https://www.researchgate.net/publication/333855959>
- Tene, T., Vique López, D. F., Valverde Aguirre, P. E., Orna Puente, L. M., & Vacacela Gomez, C. (2024). Virtual reality and augmented reality in medical education: An umbrella review. *Frontiers in Digital Health*, 6, 1365345. <https://doi.org/10.3389/fdghth.2024.1365345>
- Urbina López, M., Endara Estévez, M., Toapanta Mendoza, A., Guaras Pinango, M., & Quinchiguango Jitala, J. (2024). El uso de realidad aumentada en la enseñanza de ciencias naturales en educación básica. *Revista Científica Retos de La Ciencia*, 1(4), 224–238. <https://doi.org/10.53877/rc.8.19e.202409.18>
- Urresta-Yépez, R. F. (2024). Realidad aumentada en educación superior y el análisis de sus beneficios y desafíos. *Horizon Nexus Journal*, 2(2), 57–70. <https://doi.org/10.70881/hnj/v2/n2/39>
- Vera Mora, G. R., Sanz, C., Baldassarri, S., & Coma, T. (2023). Entornos virtuales de enseñanza y aprendizaje gamificados a la luz del concepto de presencia: Revisión sistemática de literatura. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, 33, e3. <https://doi.org/10.24215/18509959.33.e3>
- Volioti, C., Orovas, C., Sapounidis, T., Trachanas, G., & Keramopoulos, E. (2023). Augmented reality in primary education: An active learning approach in mathematics. *Computers*, 12(10), 207. <https://doi.org/10.3390/computers12100207>
- Voreopoulou, A., Mystakidis, S., & Tsinakis, A. (2024). Augmented reality escape classroom game for deep and meaningful English language learning. *Computers*, 13(1), 24. <https://doi.org/10.3390/computers13010024>
- Yaniawati, P., Sudirman, Mellawaty, Indrawan, R., & Mubarika, M. P. (2023). The potential of mobile augmented reality as a didactic and pedagogical source in learning geometry 3D. *Journal of Technology and Science Education*, 13(1), 4–22. <https://doi.org/10.3926/jotse.1661>
- Zhou, Z., Oveissi, F., & Langrish, T. (2024). Applications of augmented reality (AR) in chemical engineering education: Virtual laboratory work demonstration to digital twin development. *Computers & Chemical Engineering*, 188, 108784. <https://doi.org/10.1016/j.compchemeng.2024.108784>



## Diseño Centrado en el Usuario en Aplicaciones Web

## User-Centered Design in Web Applications

### Autores

\* **Miriam Lizeth Lectong Anchundia**   
✉ miriam\_lectong\_mdw@espm.edu.ec

**Jéssica Johanna Morales Carrillo**   
✉ jmorales@espm.edu.ec

Escuela Superior Politécnica  
Agropecuaria de Manabí Manuel  
Félix López, Dirección de Posgrado y  
Educación Continua,  
Maestría en Diseño Web y Desarrollo de  
Apps. Calceta, Manabí, Ecuador.

\*Autor para correspondencia

### Comó citar el artículo:

Lectong Anchundia, M.L. & Morales Carrillo, J.J. 2025. Diseño Basado en el Usuario en Aplicaciones Web. *Informática y Sistemas*. 9 (2), pp. 140-151. <https://doi.org/10.33936/isrtic.v9i2.7535>

Enviado: 19/05/2025  
Aceptado: 22/08/2025  
Publicado: 26/08/2025

### Resumen

El Diseño Centrado en el Usuario se ha consolidado como un enfoque metodológico protagonista en el desarrollo de aplicaciones web, puesto que promueve la participación activa de los usuarios durante todo el ciclo de diseño. Con el objetivo de analizar el impacto de este diseño en la experiencia del usuario en aplicaciones web, se presenta esta revisión sistemática de estudios desarrollados entre 2021 y 2025, en la que se aplicó la metodología de muestreo estandarizado para revisiones sistemáticas de la literatura, organizando así el proceso en cuatro fases: (1) criterios de elegibilidad, (2) búsqueda de literatura, (3) revisión basada en resúmenes y (4) lectura de textos completos. Inicialmente se definieron criterios de búsqueda, contenido y exclusión para luego realizar la búsqueda de literatura en bases de datos científicas, utilizando los términos y cadena de búsqueda; los resúmenes de los estudios recuperados se evaluaron con bases a tres preguntas clave para determinar su pertinencia y relevancia con la investigación; finalmente, se seleccionaron 20 artículos que cumplieron con todos los criterios establecidos y de los cuales se analizó el contenido completo, organizando la información relevante de cada estudio en una tabla estructurada. Los resultados demuestran una mejora notable en la experiencia de los usuarios tras el uso del diseño centrado en el usuario, integrando interfaces más intuitivas, eficientes y accesibles, desarrolladas en torno a sus expectativas y necesidades; además, refuerzan los beneficios de este enfoque en el desarrollo de aplicaciones centradas en las personas, más sostenibles.

**Palabras clave:** Diseño centrado en el usuario; experiencia del usuario; usabilidad; accesibilidad; aplicaciones web.

### Abstract

User-Centered Design has become a key methodological approach in the development of web applications, as it fosters the active participation of users throughout the entire design cycle. Aiming to analyze the impact of this design on user experience in web applications, this systematic review examines studies conducted between 2021 and 2025, following the standardized sampling for systematic literature reviews methodology. The review process was organized into four phases: (1) eligibility criteria, (2) literature search, (3) abstract-based screening, and (4) full-text reading. Initially, specific search, content, and exclusion criteria were established, followed by a systematic literature search across scientific databases using defined search terms and query strings. The abstracts of the retrieved studies were assessed based on three key questions to determine their pertinence and relevance to the research objective. Finally, 20 articles that met all established criteria were selected, and their full content was analyzed, with the relevant information from each study organized into a structured table. The results reveal a notable improvement in user experience following the implementation of user-centered design, with the development of more intuitive, efficient, and accessible interfaces tailored to users' expectations and needs. Furthermore, the findings reinforce the benefits of this approach in creating more sustainable, people-centered web applications.

**Keywords:** User-Centered Design; User Experience; Usability; Accessibility; Web Applications.



## 1. Introducción

El Diseño Centrado en el Usuario (UCD) se ha convertido en un enfoque clave en el desarrollo de aplicaciones web que promueve la creación de interfaces priorizando las necesidades, expectativas y experiencias de los usuarios, ubicándolos como eje principal de su diseño y desarrollo. Investigaciones recientes demuestran que la aplicación del UCD favorece de manera significativa a la reducción de brechas de accesibilidad en la interacción humano-computadora, puesto que incrementa la usabilidad y la satisfacción del usuario (He et al., 2023; Poduval et al., 2022).

La usabilidad es uno de los aspectos fundamentales del UCD. Turesson et al. (2022) y Herlambang et al. (2024), en sus estudios, destacan que el diseño adecuado de una interfaz facilita el acceso y la navegación; además, minimiza la carga cognitiva de los usuarios al realizar y completar mayor número de tareas en menor tiempo. Adicional a esto, Mena Mena et al. (2025), con respecto al aprendizaje en entornos digitales, sostienen que un diseño centrado en el usuario contribuye a la mejora del proceso al permitir la retroalimentación inmediata con interfaces accesibles que facilitan la navegación. Otro aspecto relevante dentro de este enfoque es la accesibilidad, con la que se garantiza el acceso sin barreras a personas con diversas capacidades, de manera equitativa (Chinchay et al., 2024).

En la actualidad, el éxito de una aplicación web se mide no solo por su capacidad técnica y operativa, sino también por la experiencia que ofrece al usuario durante la interacción. A pesar de ello, son muchas las plataformas que fracasan precisamente por no comprender y no lograr satisfacer las necesidades de sus usuarios y, sobre todo, no incluir al usuario en el diseño y desarrollo, lo que da como resultado interfaces poco o nada amigables que complican la navegación y generan frustración en quienes hacen uso de la aplicación, situación que en muchas ocasiones termina en el abandono del aplicativo por parte de los usuarios. Esta situación, según Saunders et al., (2022), se genera debido a la implementación superficial, incompleta u omisión del UCD en el ciclo de desarrollo de los productos de software impidiendo, por otra parte, la adopción y producción de soluciones digitales eficaces y sostenibles.

En cuanto a la evaluación del UCD, existen diversas métricas estandarizadas que permiten cuantificar objetivamente la experiencia del usuario y, de esta manera, facilitar la comparación entre estudios. Entre las más utilizadas, y que están presente en varios de los estudios analizados en esta revisión, se encuentran el System Usability Scale (SUS), que evalúa el nivel de usabilidad percibida con una puntuación entre 0 y 100; el User Experience Questionnaire (UEQ), que mide características específicas como

lo atractivo, eficiente y claro de la aplicación; y el Importance-Performance Analysis (IPA), que relaciona la relevancia percibida de las características, por parte de los usuarios, con su rendimiento real.

En definitiva, diseños que adoptan enfoques centrados en el usuario logran experiencias más significativas, lo cual influye directamente en la lealtad, retención y satisfacción a largo plazo. Por ello, el UCD se consolida no solo como una metodología de diseño, sino como una herramienta estratégica para alinear los productos digitales con las expectativas reales de quienes los utilizan (Othman et al., 2024; Frantzvaag et al., 2025).

En atención a estas premisas, el estudio que se presenta en este documento tiene como objetivo principal analizar el impacto del UCD en la experiencia del usuario (UX) en aplicaciones web. Si bien se trata de una revisión sistemática de literatura, esta investigación abarca la síntesis y análisis crítico de estudios empíricos recientes que implementaron el enfoque UCD en aplicaciones web en entornos reales; por tanto, el estudio ofrece un panorama integral que aporta evidencia comparativa y estructurada del impacto del UCD en la experiencia del usuario (UX) en diversos ámbitos, que contribuye a consolidar su validez metodológica y su valor estratégico en el desarrollo de aplicaciones web.

## 2. Materiales y Métodos

Este artículo de revisión se desarrolló con base en la metodología STAMP (Standardized Sampling for Systematic Literature Reviews), la cual contempla el cumplimiento de cuatro etapas fundamentales para organizar de manera adecuada el proceso de revisión: criterios de elegibilidad, búsqueda de literatura, revisión basada en resúmenes y lectura de textos completos (Roggé et al., 2024), asegurando un proceso sistemático, transparente y reproducible.

### 2.1. Criterios de elegibilidad

En esta etapa se estableció los parámetros que delimitaron las publicaciones que serían incluidas o excluidas del análisis, asegurando que la revisión se alineara con los objetivos de investigación. En este sentido, se definieron tres tipos de criterio: de búsqueda, de contenido y de exclusión, los que se detallan en la Tabla 1.

De esta manera se generó un protocolo de revisión que garantizó un proceso sistemático y recopilación de información relevante para el desarrollo de la investigación.

**Tabla 1:** Criterios de elegibilidad.  
Fuente: Las autoras

| Tipo de criterio | Descripción   |
|------------------|---|
|                  | Idioma: Inglés y español.   |
| De búsqueda      | Años de publicación: Estudios publicados entre 2021 y 2025, para garantizar la relevancia contemporánea del análisis. |
|                  | Tipo de documentos: Artículos de revistas, libros y capítulos de libros.  |
| De contenido     | Relación directa con UCD aplicado al desarrollo de aplicaciones web.  |
| De exclusión     | Inclusión de métricas clave como usabilidad, accesibilidad, satisfacción del usuario y tasas de conversión.           |

## 2.2. Búsqueda de literatura

El propósito de esta etapa fue recopilar publicaciones potencialmente relevantes mediante una búsqueda exhaustiva y sistemática basada en los criterios de elegibilidad definidos en la fase previa, a fin de obtener un conjunto inicial de publicaciones con información bibliográfica completa. La Figura 1 detalla los pasos que se siguieron para llevar a cabo la búsqueda:

Las actividades realizadas se detallan a continuación:

**Selección de bases de datos:** Se realizó la búsqueda en IEEE Xplore, Science Direct, Google Scholar, PubMed y Dialnet,

se detalla en la Tabla 2; se utilizaron operadores booleanos para combinar términos y ampliar la búsqueda:

**Documentación:** Cada búsqueda se documentó en una tabla con los campos de: fecha de consulta, base de datos utilizada y el número de resultados obtenidos.

## 2.3. Revisión basada en resúmenes

En esta etapa, denotada con las siglas ABS (Abstract-Based Screening) se revisaron los resúmenes de las publicaciones recopiladas, para determinar así su relevancia y relación con el tema de investigación. En la Tabla 3 se presentan los tres criterios

**Tabla 3:** Criterios de evaluación ABS.

Fuente: Las autoras

| Parámetro de evaluación | Criterios incluidos  |
|-------------------------|--|
| Preguntas de evaluación | <ol style="list-style-type: none"> <li>1. <math>P1 = \{\text{El resumen menciona explícitamente UCD o principios relacionados}\}</math></li> <li>2. <math>P2 = \{\text{El estudio aborda métricas de UX como accesibilidad, satisfacción del usuario o usabilidad}\}</math></li> <li>3. <math>P3 = \{\text{Se centra en el contexto de aplicaciones web}\}</math></li> </ol> |
| Puntaje binario         | 1 = Cumple el criterio;<br>0 = No cumple el criterio.  |
| Porcentaje de inclusión | Publicaciones con al menos un 70% de cumplimiento en las categorías fueron seleccionadas para la siguiente etapa   |



**Figura 1:** Proceso de búsqueda de la literatura científica.

Fuente: Las autoras

cuyas publicaciones incluyen investigaciones científicas de gran relevancia en diversas áreas del conocimiento.

**Construcción de la cadena de búsqueda:** Se definieron los términos de búsqueda, tanto en inglés como en español, tal como

**Tabla 2:** Términos de la cadena de búsqueda.

Fuente: Las autoras

| Términos principales  | Términos adicionales   |
|---|--|
| Ingles: “User-Centered Design” OR “UCD” AND “User Experience” OR “UX” AND “Web Applications”.             | “Human-Centred Design”, “Interaction Design”, “Diseño Centrado en el usuario” y “Diseño Interactivo” |
| Español: “Diseño Centrado en el Usuario” O “UCD” Y “Experiencia del Usuario” O “UX” Y “Aplicaciones Web”. |  |

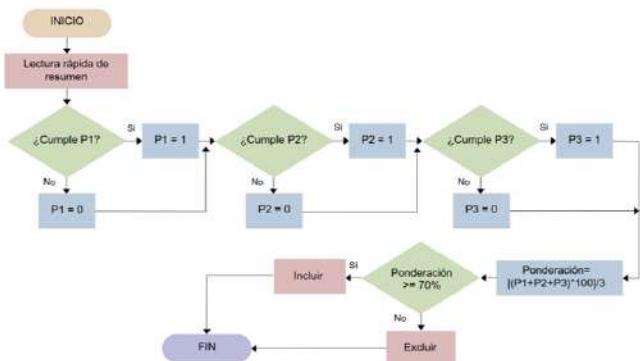
definidos, mismos que permitieron evaluar los documentos de manera objetiva.

La Figura 2 presenta de manera gráfica el flujo del proceso realizado. Inicialmente se evaluaron las preguntas de manera individual, asignándoles el valor de 1 si la respuesta era SI, caso contrario se asignaría 0. Con la sumatoria de los puntajes se realizó una regla de tres simple, obteniendo así el porcentaje de inclusión. Finalmente, si el cumplimiento de las preguntas de evaluación era igual o superior al 70%, el artículo se incluiría en la fase de lectura de textos completos.

## 2.4. Lectura de textos completos

Esta etapa, conocida también con las siglas FTR por su nombre en inglés (Full-Text Reading), se realizó la lectura completa y detallada de los textos seleccionados en la fase previa, con el objetivo de identificar patrones, métricas y relaciones clave sobre el impacto del UCD en la UX en aplicaciones web. Para ello,





**Figura 2:** Diagrama de flujo del proceso de evaluación ABS según criterios de inclusión.

Fuente: Las autoras

se diseñó una tabla en Excel con campos bien definidos, que sirvió como base para organizar y analizar los datos de manera estructurada. Los campos incluidos y su descripción se muestran en la Tabla 4:

**Tabla 4:** Campos incluidos en la tabla de análisis FTR.

Fuente: Las autoras

| Campo                 | Descripción  |
|-----------------------|--|
| ID del documento      | Identificador único asignado a cada publicación.   |
| Título                | Título de la publicación.  |
| Autor(es)             | Nombre(s) del autor(es) del estudio.   |
| Año de publicación    | Año en que fue publicado el documento.   |
| Base de datos         | Fuente de origen (Scopus, IEEE Xplore...)  |
| Implementación de UCD | Detalles sobre cómo se aplicó el UCD, identificando metodologías específicas o prácticas empleadas   |
| Métricas evaluadas    | Usabilidad (tareas completadas, tiempo de respuesta).<br>Satisfacción del usuario (Net Promoter Score, encuestas).<br>Accesibilidad, bajo cumplimiento de las WCAG 2.1 |
| Resultados clave      | Hallazgos destacados relacionados con las métricas   |

### 3. Resultados y Discusión

Con base a lo establecido en la metodología, en la primer fase se definieron los criterios de elegibilidad que se aplicaron durante la búsqueda de literatura para filtrar los resultados, bajo estos criterios se contemplaron estudios publicados en inglés y español, entre 2021 y 2025 correspondientes a artículos de revistas, libros y capítulos de libros cuyo contenido tuviera relación directa con UCD aplicado al desarrollo de aplicaciones web e incluyeran métricas como usabilidad, accesibilidad, satisfacción del usuario y tasas de conversión.

En la Tabla 5 se muestran los registros obtenidos en la búsqueda de literatura realizada en las bases de datos seleccionadas. En algunos casos se omitió el uso del operador booleano “OR” que tiene relación con la abreviatura de los términos principales: UCD o DCU y experiencia del usuario o UX. Luego de aplicar los criterios de búsqueda, términos y expansión de términos, según el caso, resultó un total de 152 artículos.

Tras aplicar los criterios de elegibilidad relacionados con el contenido y los de exclusión, los artículos que se evaluarían en

| Fecha de consulta    | Base de datos  | Resultados obtenidos |
|----------------------|----------------|----------------------|
| 20/02/2025           | IEEE Xplore    | 16                   |
| 22/02/2025           | Science Direct | 13                   |
| 06/03/2025           | PubMed         | 45                   |
| 06/03/2025           | Dialnet        | 14                   |
| 10/03/2025           | Google Scholar | 66                   |
| Resultados obtenidos |                | 152                  |

la siguiente etapa serían 35, descartando en su mayoría artículos que no estuvieron relacionados con el entorno web, otros que hacían referencia a requisitos de usuario, pero no la aplicación del UCD como tal y otros sin referencias claras de métricas relacionadas a la experiencia del usuario.

En la fase de revisión basada en resúmenes, se realizó el procedimiento detallado en la Figura 2 para cada uno de los 35 artículos resultantes de la fase anterior; lo cual dio como resultado un total de 20 artículos que cumplieron con los criterios de evaluación y que se incluyeron para el análisis en la última fase.

En la fase final, se realizó la lectura completa de los artículos

seleccionados, extrayendo datos generales e información relevante con respecto al UCD, su implementación, evaluación y resultados reportados por los autores en torno a la satisfacción del usuario. Esta información se muestra en la Tabla 6, en la que para cada artículo se presenta título y autores y año de publicación seguidos del nombre de la base de datos de donde se extrajo el estudio. En las tres últimas columnas se resume información sobre: implementación de UCD, describiendo los métodos, técnicas y/o fases del diseño aplicados; métricas evaluadas, donde se detallan las métricas relacionadas con la UX que fueron evaluadas, las normas, herramientas estandarizadas, métodos e instrumentos de evaluación utilizados para medir el impacto del UCD en la UX; y, resultados clave, en la que se resumen los hallazgos y resultados a los que llegaron los autores en sus estudios.

Para una mejor comprensión de la información presentada en la Tabla 6 sobre las métricas evaluadas, se describen brevemente

aquellas herramientas utilizadas por algunos autores en sus estudios:

- UEQ (User Experience Questionnaire): instrumento que mide seis dimensiones de la experiencia de usuario: atractivo, claridad, eficiencia, confiabilidad, estimulación y novedad, en una escala de -3 a +3.
- SUS (System Usability Scale): escala compuesta por 10 ítems, que proporciona una puntuación de usabilidad entre 0 y 100; valores superiores a 85 reflejan excelente usabilidad.
- SEQ (Single Ease Question): evaluación inmediata de la facilidad percibida para completar una tarea, usualmente medida en una escala de 1 a 7.
- IPA (Importance–Performance Analysis): método que cruza la importancia percibida de aspectos del sistema con su rendimiento, permitiendo priorizar mejoras.

**Tabla 6:** Análisis FTR.

Fuente: Las autoras

| ID | Título   | Autor(es)                         | Año  | Base de datos | Implementación de UCD   | Métricas evaluadas  | Resultados clave   |
|----|--|-----------------------------------|------|---------------|---|---|--|
| 1  | Análisis de criterios de experiencia de usuario (ux) en el desarrollo de la interfaz gráfica de usuario del módulo framework de la plataforma temonet de la Universidad de Guayaquil | ( Chicala Arroyave et al., 2021 ) | 2021 | Dialnet       | Se aplicó un enfoque basado en la UX como eje de desarrollo para una interfaz gráfica. Se siguió el modelo de “Elementos de la Experiencia del Usuario” de Garrett, alineados al paradigma del UCD.   | No se aplicaron métricas cuantitativas específicas, pero se integraron principios de <b>usabilidad (eficacia, eficiencia, compromiso, tolerancia a errores y facilidad de aprendizaje), accesibilidad</b> (según WCAG 2.0 e ISO 9241-151) y <b>experiencia emocional</b>  | El diseño UX mejoró la comprensión, accesibilidad e interacción del usuario con la plataforma. Se evidenció una mejora en la calidad de la experiencia emocional y en la facilidad de uso del sistema para usuarios con dislexia.      |
| 2  | Web-Based Structured Education for Type 2 Diabetes: Interdisciplinary User-Centered Design Approach  | (Poduval et al., 2022)            | 2022 | PubMed        | Se aplicó un enfoque interdisciplinario basado en UCD, integrando metodologías de la interacción humano-computadora (HCI) y de las ciencias de la salud. El desarrollo fue iterativo e incluyó tres fases: diseño inicial, optimización de usabilidad y pruebas en contexto real (in the wild). | Se emplearon métodos mixtos. En la etapa de prueba real se aplicaron dos cuestionarios. Además, se midieron <b>tasas de registro, de inicio y de finalización del curso</b> .   | Se reportaron mejoras en autoeficacia. El enfoque UCD mejoró la adecuación de la intervención antes de su evaluación definitiva, mostrando cómo la iteración con usuarios mejora la aceptación y potencial efectividad.                |
| 3  | Propuesta de mejora en el diseño de interfaz y experiencia de usuario (UX) en Moodle: valoración del alumnado  | (Segovia-García, 2022)            | 2022 | Dialnet       | Enfoque sistemático de rediseño centrado en principios del UCD y en normas internacionales como la ISO 9241-210. El rediseño fue implementado en Moodle con tema profesional Infinity.  | Se emplearon métricas mixtas: escala Likert para percepción general del rediseño con media $\geq 4.29$ en 0–5 y escala UEQ. Las dimensiones evaluadas obtuvieron en promedio más de 1.7 en una escala de -3 a +3. Además, se reportaron medidas de consistencia como Alfa de Cronbach ( $> 0.9$ ) y lambda de Guttman ( $> 0.93$ ). | El 76% valoró positivamente el cambio de la interfaz, destacando la sencillez y mejora en la navegación. La atracción, eficiencia y controlabilidad, fueron las mejores puntuadas, ubicándose en el 10% superior del benchmark global. |



| Nº | Título del Artículo  | (Autores, Año)                     | Fuente         | Análisis y Discusión   |   |   |
|----|--|------------------------------------|----------------|--|---|---|
|    |  |                                    |                | Resumen  | Conclusiones  |   |
| 4  | Co-Development of a Web Application (CO-VID-19 Social Site) for Long-Term Care Workers ("Something for Us"): User-Centered Design and Participatory Research Study                             | (Saunders et al., 2022)            | PubMed         | Aplicación del UCD en el diseño y desarrollo de una aplicación web.  | Mediante entrevistas se evaluó la <b>estética de la interfaz, la navegación, funcionalidad, comprensión de contenidos y confiabilidad en la información</b> . Con Google Analytics se analizó la interacción de los usuarios, los temas de mayor interés y experiencia de navegación.   | Se logró desarrollar una plataforma más ajustada a sus necesidades reales. Los usuarios valoraron la facilidad de uso, claridad en la navegación y relevancia del contenido.  |
| 5  | Development of a Digital Support Application With Evidence-Based Content for Sustainable Return to Work for Persons With Chronic Pain and Their Employers: User-Centred Agile Design Approach. | (Turesson et al., 2022)            | PubMed         | Enfoque ágil centrado en el usuario (User-Centred Agile Design) basado en cinco principios: separación de fases de descubrimiento y desarrollo, diseño iterativo con retroalimentación empírica, desarrollo paralelo, participación continua de usuarios y uso de personas/escenarios. | Se utilizaron pruebas de usabilidad con metodología "think aloud", cuestionarios de evaluación abierta, y la escala SUS. Puntuación media SUS de 95 (personas) y 88 (empleadores).  | SWEPPE fue valorado como fácil de usar, relevante y de baja carga cognitiva. El sistema logró adaptarse a las necesidades reales, mejorando la percepción de control, motivación y comunicación empleador-empleado. |
| 6  | UI/UX Design for Tourism Village Website Using the User Centered Design Method   | (Yasmine & At-mojo, 2022)          | ACM            | Ciclo completo del método UCD aplicado a un sitio web turístico, que incluyó rediseño iterativo.   | UEQ aplicado antes y después del rediseño; mejoras >2.50 en <b>atractivo, claridad, eficiencia, fiabilidad, estimulación y novedad</b> .  | Experiencia notablemente mejorada en accesibilidad y navegación; validación positiva en múltiples perfiles de usuario.  |
| 7  | User-centered design of a web-based crowdsourcing-integrated semantic text annotation tool for building a mental health knowledge base   | (He et al., 2020)                  | Science Direct | Desarrollo riguroso aplicando UCD, a través de ciclos iterativos de diseño, pruebas y retroalimentación.   | En STAT: escala SUS, con puntaje de 74/100; tiempos de tarea de 62 segundos para crowdworkers vs. 238 segundos para expertos. También se reportaron F-scores de hasta 0.84. En SWEPPE: escalas de importancia de funciones con rango 8–10/10 y se valoraron niveles de <b>motivación y utilidad</b> percibida a través de pruebas con usuarios. | En ambos casos la aplicación de UCD aumentó la adopción, redujo tiempos de desarrollo y mejoró los niveles de satisfacción y eficacia percibida por los usuarios finales.   |
| 8  | Diseño centrado en el usuario y experiencia del usuario en el sistema de control de acceso de la Universidad Libre   | (Pedraza - Gutiérrez et al., 2023) | Dialnet        | Aplicación de principios del UCD. El enfoque priorizó una interfaz intuitiva, adaptada a distintas funciones (aforo, ingreso/salida), minimizando los datos requeridos y facilitando la implementación.  | Se utilizó un test de usabilidad (no se menciona el instrumento exacto), orientado a identificar áreas de mejora en la experiencia del usuario. Se recogieron observaciones cualitativas sobre <b>facilidad de uso, comprensión de interfaz y eficiencia de navegación</b> .  | El sistema logró una interfaz sencilla para el usuario. Se reporta satisfacción parcial, con recomendaciones para mejorar interacción e intuición del sistema en actualizaciones posteriores.                       |

|    |  |                                   |                |   |   |
|----|--|-----------------------------------|----------------|---|---|
|    |  |                                   |                |   |   |
| 9  | Propuesta de diseño de una metodología de desarrollo de software basado en el enfoque UI/UX  | (Macías Vélez & Loor Navia, 2023) | Google Scholar | Se propuso la metodología RIDVP, basada en enfoques UI/UX con principios del UCD, combinando elementos de Lean UX y Agile UX.   | Encuesta a expertos en UI/UX con escalas de Likert (1 a 5) para evaluar <b>prácticas de desarrollo, satisfacción del usuario, cumplimiento de requisitos, y comprensión de la interfaz.</b> Se formularon recomendaciones prácticas que promueven interfaces accesibles que promuevan la satisfacción del usuario. Se plantea como una propuesta replicable en entornos de desarrollo profesional.  |
| 10 | The importance of Usability and User-Centered Design in Web Applications: A critical Review of Literature                                      | (Omonigho, 2023)                  | Google Scholar | Revisión sobre el papel del UCD en el desarrollo de aplicaciones web.   | No se reportan métricas evaluadas directamente, ya que el estudio es de carácter teórico. Se hace referencia a métricas comunes utilizadas en estudios revisados, como la <b>usabilidad, satisfacción del usuario, tasa de retención y costos de soporte.</b> El artículo concluye que la usabilidad y el UCD son fundamentales para el éxito de las aplicaciones web.  |
| 11 | Research, Development and Design for Product Development SAC: Interaction Design Mobile Web Application Using User Centered Design Methodology | (Andrianda & Dhewanto, 2023)      | IJCSRR         | UCD integrado con Design Thinking, que incluyó entrevistas, prototipos de alta fidelidad y pruebas en Maze.   | Tasa de éxito en ejecución de tareas del 90%, SEQ con escala de 6.42/7 y SUS con una puntuación de 87.5. Validación de claridad, eficiencia y diseño visual; experiencia altamente positiva para usuarios en entorno de servicios web.  |
| 12 | Creación de la Página Web Institucional para el Instituto Superior Tecnológico Alberto Enríquez enfocada en la Experiencia de Usuario (UX)     | (Mina Burgos et al., 2024)        | Dialnet        | Enfoque metodológico UCD, estructurado en un proceso iterativo de análisis, diseño, desarrollo y evaluación.  | Encuestas de satisfacción aplicadas a una muestra representativa de 133 personas bajo un nivel de confianza del 95%. Se evaluaron <b>satisfacción general, tiempo de permanencia, tasa de rebote y uso de funcionalidades clave.</b> Se destacó la mejora en la accesibilidad, usabilidad y experiencia general del usuario, mayor tiempo de permanencia en el sitio, reducción en la tasa de rebote y mayor interacción con funcionalidades clave.                           |
| 13 | Aplicando Design Thinking para Potenciar la Experiencia de Usuario en la Interacción con un Sitio Web Universitario                            | (Carballo et al., 2024)           | Dialnet        | Enfoque de UCD integrado con la metodología Design Thinking. El proceso combinó fases UX (análisis, diseño, implementación y evaluación) con etapas de Design Thinking (empatizar, definir, idear, prototipar, testear).                            | No se utilizaron métricas estándar, pero se evaluó <b>localización del contenido, claridad, facilidad de uso, consistencia y comprensión general del sitio</b> (escala Likert de 1 a 5), las que reportaron mejoras notables tras las iteraciones. Mejora de la accesibilidad, claridad de navegación y percepción general de la interfaz y la experiencia de búsqueda. Reducción de la frustración por parte de los estudiantes.   |
| 14 | Diseño de una aplicación web para administrar clientes, cortes y generar fichas hotspot con integración a routers MikroTik                     | (Aldana et al., 2024)             | Google Scholar | Metodología UCD, considerando perfiles reales de usuarios (administrador, empleado y usuario intermedio). Se siguió el ciclo iterativo de análisis de contexto, definición de requisitos, diseño con prototipos en Figma y evaluación por usuarios. | Mediante encuestas a 5 usuarios reales se midió: <b>estética visual, disposición de elementos, claridad de navegación, facilidad de uso, satisfacción general, comprensión de funciones y organización de menús</b> , obteniendo puntuaciones de 4 y 5 en todos los ítems. Los usuarios calificaron la experiencia de usuario como altamente satisfactoria y la aplicación en general como intuitiva, clara y eficaz. Además, se mejoró la comprensión de procesos complejos. |



|    |   |                |  |  |   |
|----|---|----------------|--|--|---|
| 15 | EffIESGen: A User-centric Web Application for Efficient Exam Scheduling and Collaboration in Education (Dahlan et al., 2024)            | Science Direct | Enfoque UCD desde la etapa inicial. Se aplicó metodología ágil (Scrum), con sprints, prototipos y validaciones iterativas.   | Se recopilaron opiniones mediante entrevistas y cuestionarios tras el uso del sistema en fase beta, para documentar observaciones cualitativas y escenarios de uso real que reflejan la efectividad y aceptación de la aplicación.   | La UX se valoró como positiva, con énfasis en accesibilidad, simplicidad, adaptabilidad y utilidad. El sistema demostró ser una herramienta innovadora que mejora la gestión académica y el entorno de aprendizaje.             |
| 16 | Unlocking inclusive education: A quality assessment of software design in applications for children with autism (Chinchay et al., 2024) | ACM            | Se utilizó una heurística específica para el UCD desarrollada por Gallardo-Montes, que incluye tres dimensiones: diseño/formato, contenido y pedagogía. El estudio incluyó fases de recolección, clasificación, evaluación y análisis de características clave en base a las necesidades psicopedagógicas del usuario con TEA. | Sistema de indicadores dividido en tres dimensiones: <b>Diseño/Formato</b> , 22 pts.; <b>Contenido</b> , 18 pts.; y, <b>Pedagogía</b> , 6 pts. Se aplicaron pruebas estadísticas no paramétricas como Kolmogorov-Smirnov, correlaciones con las que se exploró la influencia y el costo de la app. | El estudio concluye que tanto las AT como las mainstream deben ser mejoradas con principios UCD: interfaces simples, contenidos personalizables, integración de sistemas de comunicación diversos y actualizaciones sostenidas. |
| 17 | A Comprehensive User Experience Analysis of Cultural Heritage Progressive Web App Using A Hybrid UEQ-IPA Approach (Othman et al., 2024) | ACM            | UCD integrado con metodología ágil, mediante ciclos iterativos de validación con usuarios.   | Cuestionario UEQ + Análisis IPA, evaluando criterios como <b>atractivo, claridad, eficiencia, confianza, estimulación y novedad</b> .  | Alto rendimiento en atractivo y eficiencia; dimensiones de novedad y fiabilidad por debajo del promedio; mejora en percepción general de UX.  |
| 18 | Application Of User Centered Design (UCD) Method For UI/UX Design At Husqy Petshop (Herlambang Cahya Pratama et al., 2024)              | MSJ            | UCD aplicado en todas sus fases con diseño iterativo y validación para interfaces web y móviles del sistema de atención veterinaria.   | Uso de SEQ con puntaje promedio obtenido de 6.89/7 en tareas clave para cada tipo de usuario.  | Interfaz mejorada para usuarios y administradores; navegación clara, personalización y funcionalidad efectiva valoradas por usuarios.   |
| 19 | Diseño y Usabilidad de un Sistema Web para Fortalecer el Aprendizaje Autónomo en la Educación Superior (Franco Cantos et al., 2025)     | Dialnet        | Enfoque de UCD en combinación con estrategias de aprendizaje autónomo. Uso de la técnica JAD en la definición de criterios e implementación de un proceso iterativo.   | Ejecución de tareas específicas dentro del rango aceptable, con margen de error de $\pm 0.30$ segundos respecto al tiempo óptimo. 93,3% de <b>aceptación en el diseño</b> ; 95,7% en <b>usabilidad</b> ; y, 94,3% en <b>contenido</b> .  | El diseño del sistema se alineó con las necesidades reales, mejorando la efectividad en el proceso de aprendizaje autónomo.   |

- 20 MusicReco: Interactive Interface Modelling with User-Centered Design in a Music Recommendation System (Frantzvaag et al., 2025) IEEE

Enfoque iterativo basado en el UCD, estructurado en dos talleres con usuarios finales y expertos y se aplicaron principios de Design Thinking. Se aplicaron escalas como la TRL en nivel 4 (validación en laboratorio), tipo Likert, análisis de tareas y evaluaciones heurísticas. Además, coeficientes de confiabilidad como Cronbach's Alpha (valores promedio  $>0.70$ ) y McDonald's Omega ( $\omega \approx 0.92$ ). La aplicación fue altamente valorada por los usuarios, mostrando una experiencia más transparente y participativa respecto a sistemas de recomendación tradicionales.

Tras el análisis de la literatura seleccionada para este estudio, los resultados muestran una tendencia consistente hacia la efectividad del enfoque UCD en la UX. Se identificaron principios de UCD aplicados en sus diversas etapas, desde la recolección de requisitos hasta la evaluación de prototipos finales.

Algunos estudios evaluaron la UX mediante escalas estandarizadas, cuyos resultados se sitúan en rangos altos de aceptación. En su estudio, Turesson et al. (2022) reportaron una puntuación promedio de 95 en la escala SUS, que se traduce en un alto nivel de satisfacción y eficiencia, de acuerdo con los criterios establecidos por establecidos por Brooke (2013), donde valores superiores a 85 corresponden a este nivel; por su parte, Andrianda & Dhewanto (2023) aplicaron SEQ, con una puntuación de 6.42/7, que equivale a un nivel óptimo de facilidad de uso, y SUS con un puntaje de 87.5, lo que confirmó la claridad y eficiencia del diseño con los usuarios; asimismo, Segovia-García (2022) en los resultados de su estudio obtuvo valores superiores a 1.7 en todas las dimensiones del UEQ, situando la interfaz de la aplicación en el top 10% internacional según los baremos de Schrepp et al. (2017). Estos resultados muestran un patrón común que sugiere que cuando el diseño se desarrolla con un enfoque UCD, las evaluaciones tienden a superar los valores de referencia según los estándares, lo que se asocia con altos niveles de satisfacción y usabilidad percibida por parte de los usuarios.

En contraste con los estudios anteriores, otros aplicaron opciones más flexibles y contextuales, como encuestas adaptadas, validaciones participativas y escalas simplificadas que, si bien no cuentan con baremos internacionales, si permiten captar la percepción de los usuarios en entornos reales de manera directa. En este contexto, Mina Burgos et al. (2024), mediante encuestas de satisfacción post-uso, registraron un incremento en la satisfacción, tiempo de permanencia y uso de funcionalidades clave; Dahlán et al. (2024), por su parte, aplicaron entrevistas y cuestionarios que, tras el uso del sistema en fase beta, evidenciaron mejoras en dimensiones como accesibilidad, adaptabilidad y utilidad percibida por parte de los usuarios; asimismo, Aldana et al. (2024) reportaron puntajes de 4 y 5 (en una escala de 1 al 5, donde 5 representa total satisfacción) en facilidad de uso, claridad y satisfacción general, tras el rediseño iterativo con usuarios reales. En conjunto, los resultados presentados indican que, aun sin recurrir a instrumentos estandarizados al momento de evaluar los sistemas o aplicaciones, el involucrar al usuario

tempranamente en el proceso de desarrollo puede generar mejoras tangibles en experiencia percibida.

Estudios como los de Franco Cantos et al. (2025), Carballo et al. (2024) y Chinchay et al. (2024), aplicados en contextos educativos, muestran evidencia de mejoras específicas en cuanto a aprendizaje autónomo, reducción de frustración por parte de los estudiantes y otros usuarios y accesibilidad cognitiva, respectivamente, reflejando la manera en que los estudiantes se benefician de interfaces adaptadas a su proceso de aprendizaje. En el ámbito de la salud, Poduval et al. (2022) y Saunders et al. (2022) destacaron una mejor percepción de autoeficacia, inclusión y utilidad de las plataformas, lo que revela que los usuarios valoran en gran medida la confianza y seguridad en sus propias capacidades con respecto al manejo de información relacionada con su bienestar. De la misma manera, Yasmine & Atmojo (2022) y Othman et al. (2024) mostraron en sus estudios mejoras cuantificables superiores a 2.5 puntos en las dimensiones de atractivo, fiabilidad, estimulación y claridad, usando UEQ e IPA en entornos turísticos y culturales, lo que denota la importancia de diseñar interfaces claras, confiables y visualmente funcionales que motiven la interacción y exploración en experiencias que dependen del descubrimiento y la planificación. Estos hallazgos, en términos generales, apuntan a que el UCD no solo mejora la UX, sino que permite adaptar soluciones digitales a las prioridades y expectativas de los usuarios en diversos dominios, de manera eficiente.

En conjunto, los hallazgos evidencian que el enfoque UCD no solo mejora los aspectos funcionales del sistema, sino que potencia la percepción subjetiva del usuario. Estos, a su vez, coinciden con lo propuesto por Gao et al. (2021), quienes sostienen que un diseño centrado en el usuario no solo mejora la funcionalidad del sistema, sino que contribuye también en su aceptación y eficacia en contextos reales. Asimismo, en lo referente a las métricas de evaluación, Garrido et al. (2022) exponen que herramientas como SUS y UEQ permiten valorar la percepción del usuario con una alta confiabilidad, lo que se confirma en aquellos estudios en los que se aplicaron estas métricas y se obtuvieron resultados positivos que superaron los estándares de referencia.

De los 20 artículos analizados, 17 reportaron mejoras tangibles en usabilidad, satisfacción, accesibilidad o eficiencia tras aplicar principios UCD; no obstante, a pesar de la consistencia en los beneficios reportados, algunos estudios presentan limitaciones



metodológicas importantes. Por ejemplo, estudios como los desarrollados por Chicala Arroyave et al. (2021) o Carballo et al. (2024) carecen de métricas cuantitativas o reportan resultados sin escalas estandarizadas, lo que dificulta la comparación transversal entre investigaciones, además, no todos los estudios documentan el tamaño de muestra, las características sociodemográficas de los participantes ni el número de iteraciones realizadas en el proceso de diseño. Por otra parte, algunos artículos presentan sesgos de implementación; por ejemplo, Saunders et al. (2022) destacan una baja tasa de finalización a pesar de haber seguido un enfoque participativo, lo cual evidencia que factores como la motivación o el contexto sociocultural también influyen en la experiencia final del usuario. Esto refuerza lo mencionado por Bourazeri y Stumpf (2021), quienes señalan que el éxito del UCD requiere también una adecuada contextualización social y operativa del producto digital.

Los resultados obtenidos en esta revisión sistemática sugieren que el enfoque UCD presenta una aplicación cada vez más frecuente y efectiva en el desarrollo de aplicaciones web, y se perfila como una metodología clave para mejorar la UX en diversos contextos.

#### 4. Conclusiones

En relación con el objetivo de esta revisión sistemática, los resultados obtenidos a partir del análisis de 20 estudios publicados en diversos contextos sostienen, de manera consistente, que la aplicación del UCD se asocia estrechamente con mejoras en aspectos relacionados con la UX, como eficiencia, claridad de navegación, estética de la interfaz y satisfacción general del usuario. Sin embargo, la interpretación de estas conclusiones debe considerar la heterogeneidad de metodologías, métricas y dominios de la muestra de este estudio.

Los resultados de esta revisión indican que el UCD no solo mejora la usabilidad y otros criterios de la UX desde la perspectiva del usuario, sino que puede influir de manera positiva en la aceptación y sostenibilidad de las aplicaciones web. En este sentido, estos hallazgos pueden servir como base para que los equipos de diseño prioricen la integración de metodologías participativas y el uso de métricas estandarizadas, como SUS, UEQ o IPA, durante el ciclo de desarrollo iterativo de aplicaciones web, lo que ayudaría a una mayor alineación de estas soluciones digitales con las necesidades reales de los, haciéndolas más útiles, pertinentes y sostenibles.

Futuras investigaciones pueden explorar el impacto del UCD desde enfoques longitudinales, así como también su aplicación en tecnologías emergentes como la inteligencia artificial, donde la interacción humano-máquina necesita nuevos enfoques que

garanticen experiencias satisfactorias.

Esta revisión presenta limitaciones tanto propias como derivadas de los estudios analizados. En cuanto a las propias del estudio, al limitar la búsqueda a publicaciones en inglés y español se pudieron haber omitido estudios de gran relevancia escritos y/o publicados en otros idiomas; se incluyó un total de 20 artículos, lo que puede considerarse una muestra reducida y que podría influir en la amplitud y solidez de las conclusiones; no se realizó un metaanálisis estadístico de los resultados, debido a la heterogeneidad de las métricas, herramientas, enfoques y ámbito de aplicación, restringiendo así la capacidad de sintetizarlos en términos cuantificables y comparables. Respecto a las derivadas, algunas de las limitaciones metodológicas que se encontraron tienen que ver con la muestra de estudio, que en varios casos fueron muy reducidas; así también, la variabilidad de instrumentos de evaluación aplicados, algunos bajo escalas estandarizadas y otros usando instrumentos más cualitativos y flexibles y la diversidad de contextos de las aplicaciones, lo que podría afectar la generalización de los hallazgos.

#### Contribución de los autores

**Miriam Lizeth Lectong Anchundia:** Investigación, Redacción - borrador original del artículo. **Jéssica Johanna Morales Carrillo:** Metodología, Redacción - revisión y edición del artículo.

#### Conflictos de interés

Los autores no tienen conflictos de intereses.

#### Referencias bibliográficas

- Aldana, A. C., Rodríguez, C. J. S., Luna, A. H., & Cabrera, H. H. (2024). Diseño de una aplicación web para administrar clientes, cortes y generar fichas hotspot con integración a routers MikroTik. *Revista de Investigación en Tecnologías de la Información*, 12(26), 1–18. <https://doi.org/10.36825/RITI.12.26.001>
- Andriana, J. M., & Dhewanto, W. (2023). Research, development and design for product development SAC: Interaction design mobile web application using user-centered design methodology. *International Journal of Current Science Research and Review*, 6(12), 8021–8031. <https://doi.org/10.47191/IJCSRR/V6-I12-82>
- Brooke, J. (2013). SUS: A retrospective. *Journal of Usability Studies*, 8(2), 29–40. <https://doi.org/10.5555/2817912.2817913>

- Carballo, M. L., Lopez, M. E., & Martin, A. (2024). Aplicando design thinking para potenciar la experiencia de usuario en la interacción con un sitio web universitario. *Informes Científicos Técnicos - UNPA*, 16(1), 71–98. <https://doi.org/10.22305/ict-unpa.v16.n1.1100>
- Chicala Arroyave, J., Arizaga Gamboa, J., & Alvarado Unamuno, E. (2021). Análisis de criterios de experiencia de usuario (UX) en el desarrollo de la interfaz gráfica de usuario del módulo framework de la plataforma Temonet de la Universidad de Guayaquil. Serie Científica de la Universidad de las Ciencias Informáticas, 14(9), 95–105. <https://dialnet.unirioja.es/servlet/articulo?codigo=8590644>
- Chinchay, Y., Gomez, J., & Montoro, G. (2024). Unlocking inclusive education: A quality assessment of software design in applications for children with autism. *Journal of Systems and Software*, 217, Article 112164. <https://doi.org/10.1016/j.jss.2024.112164>
- Dahlan, M. M., Jantan, S. R., Mohtaram, S., & Azmi, M. A. A. (2024). EffiESGen: A user-centric web application for efficient exam scheduling and collaboration in education. *Procedia Computer Science*, 245, 1092–1100. <https://doi.org/10.1016/j.procs.2024.10.338>
- Franco Cantos, J. H., Zambrano Zambrano, D. M., Vera Pico, L. A., & Zambrano Acosta, J. M. (2025). Diseño y usabilidad de un sistema web para fortalecer el aprendizaje autónomo en la educación superior. *Reincisol*, 4(7), 987–1006. [https://doi.org/10.59282/REINCISOL.V4\(7\)987-1006](https://doi.org/10.59282/REINCISOL.V4(7)987-1006)
- Frantzvaag, M., Chatterjee, A., Ghose, D., & Dash, S. P. (2025). MusicReco: Interactive interface modelling with user-centered design in a music recommendation system. *IEEE Access*, 13, 22345–22358. <https://doi.org/10.1109/ACCESS.2025.3540201>
- He, X., Zhang, H., & Bian, J. (2020). User-centered design of a web-based crowdsourcing-integrated semantic text annotation tool for building a mental health knowledge base. *Journal of Biomedical Informatics*, 110, Article 103571. <https://doi.org/10.1016/j.jbi.2020.103571>
- Herlambang Cahya Pratama, Y., Al Hafidz, M., Lazuardy, N., & Naristi, K. (2024). Application of user centered design (UCD) method for UI/UX design at Husqy Petshop. *MSJ: Majority Science Journal*, 2(2), 62–70. <https://doi.org/10.61942/msj.v2i2.152>
- Macías Vélez, E. M., & Loor Navia, E. A. (2023). Propuesta de diseño de una metodología de desarrollo de software basado en el enfoque UI/UX. *Código Científico Revista de Investigación*, 4(2), 387–405. <https://doi.org/10.55813/gaea/ccri/v4/ne2/167>
- Mena Mena, C. C., Castillo Duchi, A. M., Fernández Acevedo, J. E., & Jiménez Granizo, C. N. (2025). Diseño y usabilidad de una app para la enseñanza de las fases del ciclo celular. *Informática y Sistemas Revista de Tecnologías de la Informática y las Comunicaciones*, 9(1), 1–15. <https://doi.org/10.33936/isrtic.v9i1.7014>
- Mina Burgos, R. M., Nazareno Cabezas, G. F., & Nazareno Cabezas, M. D. (2024). Creación de la página web institucional para el Instituto Superior Tecnológico Alberto Enríquez enfocada en la experiencia de usuario (UX). *Reincisol*, 3(6), 3018–3038. [https://doi.org/10.59282/reincisol.V3\(6\)3018-3038](https://doi.org/10.59282/reincisol.V3(6)3018-3038)
- Omonigho, T. J. (2023). The importance of usability and user-centered design in web applications: A critical review of literature [Manuscript]. ResearchGate. <https://doi.org/10.13140/RG.2.2.19014.04160>
- Othman, M. K., Norman Anuar, N., Barawi, M. H., Yahya, A. S. A. H., & Abdul Manaf, A. A. (2024). A comprehensive user experience analysis of cultural heritage progressive web app using a hybrid UEQ-IPA approach. *Journal on Computing and Cultural Heritage*, 17(3), Article 45. <https://doi.org/10.1145/3647998>
- Pedraza-Gutiérrez, S. I., Romero-González, J. F., Güiza-Rodríguez, J. C., & Giraldo-Henao, E. W. (2023). Diseño centrado en el usuario y experiencia de usuario en el sistema de control de acceso de la Universidad Libre. *Revista Científica de Sistemas e Informática*, 3(1), Article e426. <https://doi.org/10.51252/rksi.v3i1.426>
- Poduval, S., Ross, J., Pal, K., Newhouse, N., Hamilton, F., & Murray, E. (2022). Web-based structured education for type 2 diabetes: Interdisciplinary user-centered design approach. *JMIR Human Factors*, 9(1), Article e31567. <https://doi.org/10.2196/31567>
- Rogge, A., Anter, L., Kunze, D., Pomsel, K., & Willenbrock, G. (2024). Standardized sampling for systematic literature reviews (STAMP method): Ensuring reproducibility and replicability. *Media and Communication*, 12, Article 7836. <https://doi.org/10.17645/mac.7836>
- Saunders, C. H., Sierpe, A., Stevens, G., Elwyn, G., Cantrell, M., Engel, J., Gonzalez, M., Hayward, M., Huebner, J., Johnson, L., Jimenez, A., Little, N. R., McKenna, C., Onteeru, M., Khine, M. O., Pogue, J., Vargas, J. L. S., Schmidt, P., Thomeer, R., & Durand, M. A. (2022). Co-development of a web application (COVID-19 Social Site) for long-term care workers (“Something for Us”): User-centered design and participatory research study. *Journal of Medical Internet Research*, 24(6), Article e38359. <https://doi.org/10.2196/38359>
- Schrepp, M., Thomaschewski, J., & Hinderks, A. (2017). Design and evaluation of a short version of the user experience questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(6), 103–108. <https://doi.org/10.9781/ijimai.2017.09.001>



Segovia-García, N. (2022). Proposal to improve the interface design and user experience (UX) in Moodle: Student feedback. *EduTEC*, 82, 199–216. <https://doi.org/10.21556/edutec.2022.82.2673>

Turesson, C., Liedberg, G., & Björk, M. (2022). Development of a digital support application with evidence-based content for sustainable return to work for persons with chronic

pain and their employers: User-centered agile design approach. *JMIR Human Factors*, 9(1), Article e33571. <https://doi.org/10.2196/33571>

Yasmine, H. T., & Atmojo, W. T. (2022). UI/UX design for tourism village website using the user centered design method. *TIERS Information Technology Journal*, 3(2), 100–114. <https://doi.org/10.38043/tiers.v3i2.3871>



## Buenas Prácticas de Seguridad para la Implementación de Inteligencia Artificial en Entornos de Computación en la Nube

### Best Practices for Securing the Implementation of Artificial Intelligence in Cloud-Based Environments

#### Autores

\* **Yasbeck Jemima Mora Chávez**   
✉ [ymora6@utmachala.edu.ec](mailto:ymora6@utmachala.edu.ec)

**Reinaldo Benedicto Fernández Gonzalez**   
✉ [rfernande4@utmachala.edu.ec](mailto:rfernande4@utmachala.edu.ec)

**Joofre Antonio Honores Tapia**   
✉ [jhonores@utmachala.edu.ec](mailto:jhonores@utmachala.edu.ec)

**Milton Rafael Valarezo Pardo**   
✉ [mvalarezo@utmachala.edu.ec](mailto:mvalarezo@utmachala.edu.ec)

#### Resumen

La incorporación de la Inteligencia Artificial (IA) en los entornos de Cloud Computing (Computación en la Nube) ha permitido alcanzar mayores niveles de escalabilidad, eficiencia y automatización en la creación de sistemas inteligentes. Sin embargo, esta convergencia también ha resaltado nuevas amenazas y desafíos críticos en aspectos de seguridad y privacidad. El presente artículo desarrolla una revisión sistemática de literatura científica publicada entre 2021 y 2025, siguiendo la metodología PRISMA 2020, para identificar buenas prácticas de seguridad aplicadas al despliegue de proyectos de IA en entornos de Computación en la Nube. Se analizaron veinticinco estudios relevantes, lo que permitió sintetizar un conjunto de recomendaciones centradas en la gestión de accesos, protección de APIs, cifrado de datos y auditoría de eventos. Se estableció un prototipo de inteligencia artificial que incorporó controles de seguridad y privacidad de datos para ser finalmente desplegado en un entorno de Computación en la Nube, buscando validar la aplicabilidad de las buenas prácticas identificadas en la revisión sistemática. Los resultados obtenidos evidencian la efectividad de los controles aplicados, así como la necesidad de adoptar estrategias proactivas de seguridad desde las fases iniciales de los proyectos de despliegue de la IA en la nube.

**Palabras clave:** inteligencia artificial; computación en la nube; seguridad informática; buenas prácticas; revisión sistemática.

#### Abstract

The incorporation of Artificial Intelligence (AI) into Cloud Computing environments has enabled greater levels of scalability, efficiency, and automation in the creation of intelligent systems. However, this convergence has also highlighted new threats and critical challenges in terms of security and privacy. The present article develops a systematic review of scientific literature published between 2021 and 2025, following the PRISMA 2020 methodology, to identify good security practices applied to the deployment of AI projects in cloud computing environments. Twenty-five relevant studies were analyzed, which allowed for the synthesis of a set of recommendations focused on access management, API protection, data encryption, and event auditing. An artificial intelligence prototype incorporating security and data privacy controls was established and finally deployed in a cloud computing environment to validate the applicability of the best practices identified in the systematic review. The results obtained demonstrate the effectiveness of the controls applied, as well as the need to adopt proactive security strategies from the initial phases of AI deployment projects in the cloud.

**Keywords:** artificial intelligence; cloud computing; information security; best practices; systematic review.

Universidad Técnica de Machala,  
Facultad de Ingeniería Civil, Carrera de  
Tecnologías de la Información, Machala,  
El Oro, Ecuador.

\*Autor para correspondencia

#### Comó citar el artículo:

Mora Chavez, Y.J., Fernández Gonzalez, R.B., Honores Tapia, J.A. & Valarezo Pardo, M.R. 2025. Buenas Prácticas de Seguridad para la Implementación de Inteligencia Artificial en Entornos de Computación en la Nube. *Informática y Sistemas*. 9 (2), pp. 152-163. <https://doi.org/10.33936/isrtic.v9i2.7687>

Enviado: 03/07/2025

Aceptado: 19/08/2025

Publicado: 26/08/2025



## 1. Introducción

El desarrollo acelerado de tecnologías como la IA junto a la Computación en la Nube ha cambiado de forma radical el diseño y funcionamiento de los sistemas digitales en diferentes áreas. Por un lado, la IA automatiza procesos, permite analizar grandes volúmenes de datos y facilita la toma de decisiones; por otro, la computación en la nube brinda una infraestructura flexible, escalable y accesible para desplegar estas soluciones sin requerir recursos locales (Chen & Ying, 2022; Sunyaev, 2020; Bello et al., 2021). Esta combinación ha contribuido al aumento de los modelos de IA como servicio (AIaaS) para aplicaciones inteligentes en sectores como: salud, comercio electrónico, educación y seguridad pública (Abioye et al., 2021; Zhang et al., 2024; Kyivska & Tsutsui, 2021). A su vez, esta convergencia ha dado lugar a arquitecturas distribuidas que dan soporte a operaciones en tiempo real, integran datos de diversas fuentes y se adaptan rápidamente a las variaciones del entorno. Por el contrario, este avance conlleva nuevos retos en términos de la seguridad de estos sistemas. Varios estudios han identificado amenazas recurrentes como el envenenamiento de datos, ataques adversarios, filtración de información sensible, manipulación de APIs y una gestión inefficiente del control de accesos (Calle & Barriga, 2025; Bai et al., 2025). Estas vulnerabilidades aumentan su gravedad en entornos multiusuario, donde los datos se almacenan de forma compartida y las interfaces abiertas, como las APIs, son puntos críticos de exposición. Si bien existen marcos normativos reconocidos, como la norma ISO/IEC 27001, el OWASP Top 10 for Large Language Models (OWASP, 2023) y el NIST AI Risk Management Framework (NIST, 2023), estos proporcionan guías generales que requieren ser adaptadas a un contexto determinado de IA en la nube. Por ejemplo, la mayoría de estos no contemplan de forma específica temas como el control del aprendizaje continuo, la integridad de los conjuntos de entrenamiento, la auditabilidad de decisiones automatizadas y la supervisión post-despliegue de modelos. Esto genera una diferencia significativa entre las recomendaciones normativas y la realidad operativa en entornos distribuidos basados en IA (Grechi et al., 2025; Flores-Cedeño et al., 2024; Gutiérrez Rodríguez & Castellanos-Sánchez, 2023; Acosta Cortez, 2024). Pese a que muchos estudios abordan los temas de seguridad de la IA o de la nube, suelen hacerlo por separado, sin proponer soluciones conjuntas o validar su aplicabilidad mediante casos experimentales. Por ejemplo, trabajos como el de Alnami et al. (2025) se enfocan puntualmente en la protección de modelos de IA frente a ataques adversarios, sin contemplar los riesgos asociados al entorno de despliegue en la nube. A diferencia del estudio mencionado, Bai et al. (2025) analizan vulnerabilidades específicas en plataformas de computación en la nube, como

la gestión de accesos y el aislamiento de recursos, pero sin considerar la integración con sistemas de IA. Por su parte, Ahmed et al. (2025) proponen un marco de cifrado para proteger datos en entornos cloud, pero con un enfoque centrado en el almacenamiento seguro, sin integrar elementos de seguridad específicos de modelos de IA. Estos estudios, aunque valiosos, no consideran de forma integral los desafíos de seguridad que surgen al combinar IA y Cloud en un mismo entorno. En contraste, el presente estudio propone un enfoque original al integrar la revisión teórica y la validación práctica ante esta convergencia, lo que representa una contribución significativa frente a trabajos anteriores que solo abordan aquellos temas de manera aislada o sin evidencia experimental.

Según el contexto, se plantea la siguiente pregunta de investigación: ¿Cómo pueden implementarse buenas prácticas de seguridad en un proyecto de IA desplegado en la nube, y qué tan favorables resultan frente a amenazas comunes? A su vez, surge la siguiente hipótesis: La implementación de buenas prácticas de seguridad permite fortalecer los sistemas de IA desplegados en la nube. El propósito de la investigación es detectar, agrupar y validar buenas prácticas de seguridad aplicables en proyectos de inteligencia artificial en la nube. Para ello, se llevó a cabo una revisión sistemática de literatura apoyada en artículos científicos del periodo 2021-2025, siguiendo los lineamientos de la metodología PRISMA 2020. Las buenas prácticas extraídas buscan mitigar vulnerabilidades críticas como por ejemplo el mencionado acceso no autorizado, la fuga de información a través de APIs, la falta de cifrado o la ausencia de registro sobre los procesos de toma de decisiones automatizadas. Como aporte, se desarrolló un prototipo basado en inteligencia artificial al que se le incorporaron algunas de las buenas prácticas de seguridad identificadas en la revisión, tales como el uso de tokens para proteger APIs, el cifrado de datos sensibles en tránsito y en reposo, entre otras. Este prototipo fue desplegado en un entorno cloud y se sometió a pruebas para evaluar la aplicabilidad de dichas prácticas.

## 2. Materiales y Métodos

Esta investigación se realizó siguiendo la metodología Preferred Reporting Items for Systematic Reviews and Meta-Analyses

(PRISMA 2020). Esta fue seleccionada por su reconocimiento internacional y por ofrecer un marco bien estructurado que permite organizar, seleccionar y evaluar críticamente la literatura científica relevante. La metodología PRISMA está estructurada en tres fases principales: Identificación, Cribado e Inclusión, ilustradas y brevemente descritas en la Figura 1:

**Figura 1.** Fases de la Metodología PRISMA 2020.

Fuente: Los Autores

## 2.1. Fase 1: Identificación

### 2.1.1. Preguntas de Investigación

Las siguientes preguntas guiaron la revisión sistemática y el desarrollo del prototipo práctico: ¿Qué buenas prácticas de seguridad propone la literatura científica reciente para proteger sistemas de inteligencia artificial desplegados en entornos de computación en la nube?, ¿Qué mecanismos, marcos o controles se han documentado para mitigar amenazas comunes en soluciones de IA en la nube? y ¿Cómo pueden implementarse estas buenas prácticas y qué tan aplicables son?

### 2.1.2. Palabras Clave y Cadenas de Búsqueda

Las palabras clave se seleccionaron considerando su frecuencia de uso en la literatura científica, así como su capacidad para recuperar estudios pertinentes que aporten evidencia sobre la mitigación de amenazas comunes en sistemas distribuidos basados en IA. Estas se complementaron con términos asociados a estándares internacionales, marcos de referencia y estrategias aplicadas (como OWASP, ISO/IEC 27001, cifrado, autenticación, etc.), con el fin de captar investigaciones que presentarán controles técnicos validados o propuestas concretas de implementación. A continuación, en la Tabla 1 se detallan las palabras clave seleccionadas, junto con su justificación. De la misma manera en la Tabla 2 se evidencia la construcción de las cadenas de búsqueda específicas en inglés y español, adaptadas a los motores de búsqueda académicos utilizados:

### 2.1.3. Selección de Bases de Datos Científicas

Para la ejecución de la búsqueda bibliográfica se seleccionaron tres bases de datos ampliamente reconocidas por su cobertura académica y rigor científico: Scopus, Google Scholar y Redalyc. Estas fuentes fueron elegidas por su diversidad de publicaciones, su acceso a literatura multidisciplinaria y la posibilidad de filtrar resultados relevantes mediante el uso de operadores booleanos, lo cual facilita una recuperación más precisa de estudios relacionados con la aplicación de la inteligencia artificial en

**Tabla 1:** Palabras clave utilizadas en la búsqueda bibliográfica.

Fuente: Los Autores

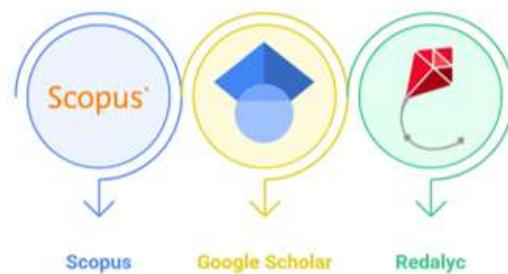
| Palabra Clave           | Justificación y contexto de uso  |
|-------------------------|--|
| Inteligencia Artificial | Concepto central del estudio. Se refiere al desarrollo de sistemas capaces de realizar tareas que normalmente requieren inteligencia humana, como el análisis de datos, toma de decisiones y automatización. |
| Cloud Computing         | Describe el entorno de operación de los sistemas de IA estudiados.   |
| Seguridad informática   | Permite enfocar la búsqueda en estudios que aborden amenazas, controles o marcos de protección.  |
| Buenas prácticas        | Palabra clave orientada a identificar recomendaciones prácticas o estrategias de implementación.   |
| Protección de datos     | Ayuda a filtrar literatura centrada en el resguardo de información personal y sensible.  |
| Marcos de seguridad     | Se refiere a estándares, normas y frameworks como OWASP, ISO/IEC 27001, NIST RMF, entre otros.   |

**Tabla 2:** Cadenas de búsqueda utilizadas en los motores de bases de datos.

Fuente: Los Autores

| Idioma  | Cadenas de Búsqueda  |
|---------|--|
| Inglés  | (Artificial Intelligence OR AI systems) AND (Cloud computing OR distributed systems) AND (security best practices OR cybersecurity frameworks OR data protection OR “security controls”)                           |
| Español | (Inteligencia Artificial OR sistemas de IA) AND (computación en la nube OR sistemas distribuidos) AND (buenas prácticas de seguridad OR marcos de ciberseguridad OR protección de datos OR controles de seguridad) |

entornos de computación en la nube. En la figura 2 se muestra las bases de datos mencionadas:

**Figura 2.** Bases de Datos Científicas utilizadas.

Fuente: Los Autores



## 2.2 Fase 2: Cribado

### 2.2.1. Criterios de Inclusión y Exclusión

Los resultados obtenidos durante la fase anterior fueron filtrados, de forma que se descartaron los documentos que no cumplían con los objetivos específicos de la investigación. El eje fundamental fue garantizar que los artículos analizados aportaran en la identificación de enunciados que enumeren buenas prácticas de la seguridad en entornos de IA, concretamente cuando se despliegan en una nube. A continuación, en la Tabla 3 se detallan los criterios definidos:

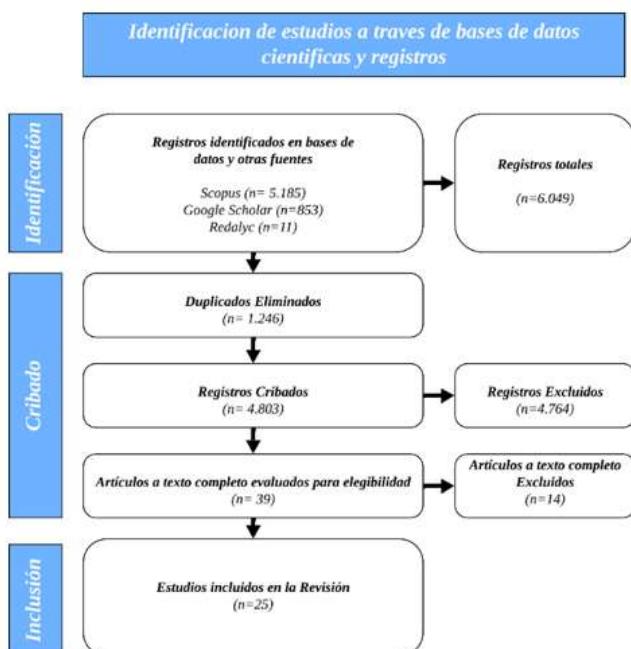
**Tabla 3:** Aplicación de criterios de inclusión y exclusión  
Fuente: Las autoras

| Tipo de criterio       | Descripción   |
|------------------------|---|
| Criterios de Inclusión | <ul style="list-style-type: none"> <li>- Artículos científicos publicados entre enero de 2021 y marzo de 2025.</li> <li>- Artículos revisados por expertos/as, y publicados en revistas indexadas o en repositorios institucionales.</li> <li>- Artículos que aborden explícitamente la seguridad de sistemas de IA desplegados en entornos de computación en la nube.</li> <li>- Artículos que propongan buenas prácticas, marcos normativos o controles técnicos aplicables a dichos entornos.</li> <li>- Artículos que estén disponibles en texto completo ya sea en idioma inglés o español.</li> </ul> |
| Criterios de Exclusión | <ul style="list-style-type: none"> <li>- Documentos previos a 2021. - Documentos centrados únicamente en IA o en el cloud computing sin abordar su interrelación.</li> <li>- Informes técnicos, TFGs o resúmenes/abstracts que carecieran de un desarrollo metodológico en su totalidad.</li> <li>- Artículos duplicados en bases de datos, o bien, documentos que no contuvieran suficiente información para su posible verificación y análisis.</li> </ul>  |

### 2.2.2. Selección de Artículos

Inicialmente se obtuvieron 6.049 registros provenientes de las bases de datos científicas mencionadas. En la fase de cribado, se eliminaron 1.246 registros duplicados. Posteriormente, al aplicar criterios de inclusión y exclusión, se descartaron 4.764 documentos por no abordar directamente temas de seguridad en sistemas de inteligencia artificial desplegados en la nube, por carecer de validez metodológica o por estar incompletos. Luego, se revisaron 39 artículos a texto completo, de los cuales 14 fueron excluidos por no presentar evidencia útil para la

formulación de buenas prácticas de seguridad. Finalmente, 25 artículos científicos fueron seleccionados para aplicarlos en la investigación ya que cumplen con los criterios establecidos, y aportan significativamente al objetivo central de la investigación. Este proceso se resume gráficamente en la Figura 3, siguiendo el diagrama de flujo PRISMA 2020:



**Figura 3.** Diagrama PRISMA 2020.

## 2.3 Fase 3: Inclusión

### 2.3.1. Evaluación de Relevancia

Se aseguró que los artículos no solo cumplieran con los criterios aplicados, sino que también fueran metodológicamente sólidos y fuera posible identificar en ellos buenas prácticas de seguridad aplicables al despliegue de sistemas de IA en la nube. Los aspectos que se evaluaron están descritos en la Tabla 4:

### 2.3.2. Extracción de Datos

Se revisó cada uno de los 25 artículos seleccionados con el fin de identificar información que pudiera contribuir a la formulación, análisis y validación de buenas prácticas de seguridad en sistemas de IA desplegados en la nube. La extracción de datos se organizó según una serie de aspectos clave, como se evidencia en la Tabla 5:

**Tabla 4:** Criterios aplicados para la evaluación de relevancia de estudios.

Fuente: Los Autores

| Aspecto Evaluado   | Detalle  |
|--|--|
| Definición del objetivo del estudio                        | Se evaluó si el artículo presenta de forma clara y específica los objetivos o preguntas de investigación.  |
| Transparencia metodológica                                 | Se verificó la descripción de la metodología utilizada, incluyendo enfoque, instrumentos o criterios de selección.                                       |
| Profundidad en el tratamiento de amenazas o riesgos        | Se analizó la profundidad con la que el estudio aborda amenazas de seguridad, prácticas de mitigación o análisis de vulnerabilidades específicas.        |
| Integración con marcos normativos o estándares             | Se consideró la incorporación de marcos como ISO/IEC 27001, OWASP, NIST RMF u otros enfoques éticos o regulatorios aplicables.                           |
| Aplicabilidad de hallazgos en contextos reales o simulados | Se valoró si las propuestas, recomendaciones o resultados del artículo son susceptibles de ser implementados en proyectos reales o prototipos simulados. |

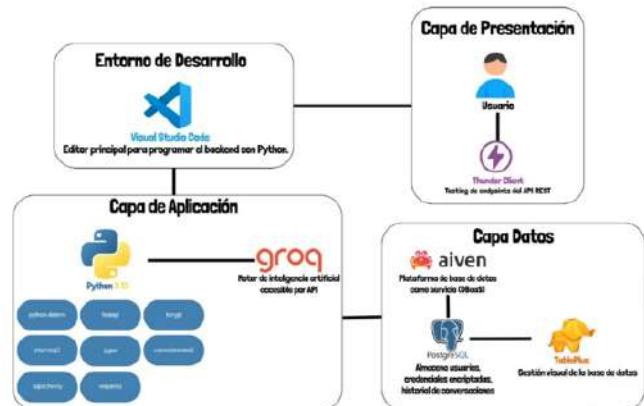
**Tabla 5:** Aspectos analizados a partir de los resultados de la extracción de datos.

Fuente: Los Autores

| Aspecto Evaluado   | Detalle   |
|--|---|
| Tipo de estudio  | Categoriza los artículos según corresponda: una revisión, estudio de caso, propuesta técnica, mapeo sistemático, entre otros. |
| Contexto de aplicación   | Describe el sector o área en el que se desarrolla la investigación.   |
| Percepciones respecto a los riesgos en el tema de la seguridad | Expone las debilidades señaladas en el estudio, como el acceso ilegal, la alteración de API, ataques maliciosos, etc.         |
| Buenas prácticas o controles sugeridos                         | Detalla acciones o recomendaciones para mitigar los riesgos mencionados.  |
| Normas o lineamientos utilizados como referencia               | Indica si el texto examinado menciona o se ajusta a estándares reconocidos como OWASP, ISO/IEC 27001, NIST u otros similares. |
| Nivel de validación o experimentación                          | Señala si las buenas prácticas fueron validadas teóricamente, mediante simulaciones, prototipos o casos reales.               |
| Aplicabilidad práctica   | Evaluá el potencial de las buenas prácticas detectadas de ser implementadas en entornos reales o simulados.                   |
| Limitaciones del estudio                                       | Identifica restricciones metodológicas o temáticas que puedan afectar la generalización de los hallazgos.                     |

## 2.4 Metodología de Desarrollo del Prototipo.

La presente investigación adopta un enfoque experimental centrado en el desarrollo de un prototipo, con el fin de validar las buenas prácticas establecidas en la revisión sistemática, fue desarrollado en el entorno de programación Visual Studio Code, usando el lenguaje Python en su versión 3.13. Para el backend, se implementó el framework FastAPI para construir una API REST segura, la autenticación fue basada en JSON Web tokens (tokens JWT) y también se empleó PostgreSQL para ser desplegado en la nube a través del servicio Aiven. Toda la información de inicio de sesión y eventos de seguridad fueron almacenados en dicha base de datos. Asimismo, el prototipo se integró con un motor de inteligencia artificial Groq, accesible por API. La fase de pruebas y validación se ejecutó utilizando la herramienta Thunder Client, permitiendo simular peticiones HTTP controladas a la API y verificar el comportamiento del sistema ante diferentes escenarios. En la Figura 4 se observa un diagrama descriptivo de la arquitectura del prototipo:

**Figura 4.** Arquitectura del prototipo.

Fuente: Los Autores

## 3. Resultados y Discusión

### 3.1 Resultados de Revisión Literaria

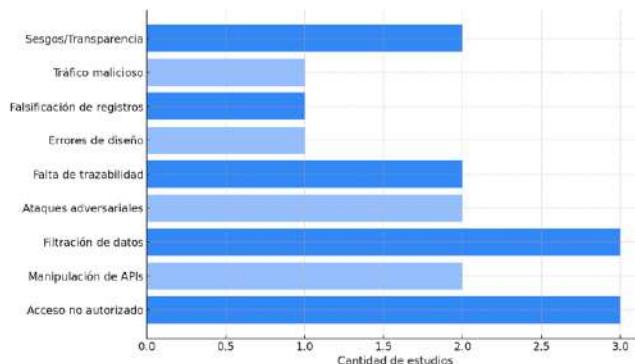
#### 3.1.1. Identificación de Amenazas Recurrentes

Los estudios analizados revelan un conjunto de amenazas comunes que afectan directamente la seguridad de los sistemas de inteligencia artificial cuando son desplegados en entornos de computación en la nube. Estas amenazas se presentan en contextos diversos como salud, redes empresariales, investigación científica, transporte inteligente y sistemas de análisis forense digital, lo que evidencia la transversalidad del problema. En la Figura 5 se ilustran las amenazas más destacadas en los artículos revisados:

#### 3.1.3. Síntesis de Buenas Prácticas extraídas

A partir del análisis de los artículos, se identificaron varias buenas prácticas de seguridad aplicables en el diseño, despliegue y operación de proyectos de inteligencia artificial en la nube:





**Figura 5.** Frecuencia de amenazas abordadas en los artículos.

Fuente: Los Autores

- Autenticación multifactor y segmentación de acceso.
- Cifrado homomórfico.
- Blockchain.
- Análisis en tiempo real de tráfico y verificación de consistencia en datos distribuidos.
- Modelado de riesgos con coanálisis de seguridad y confiabilidad (safety).
- Supervisión continua, monitoreo de eventos y registro de auditoría.
- Políticas de transparencia algorítmica y ética en IA.

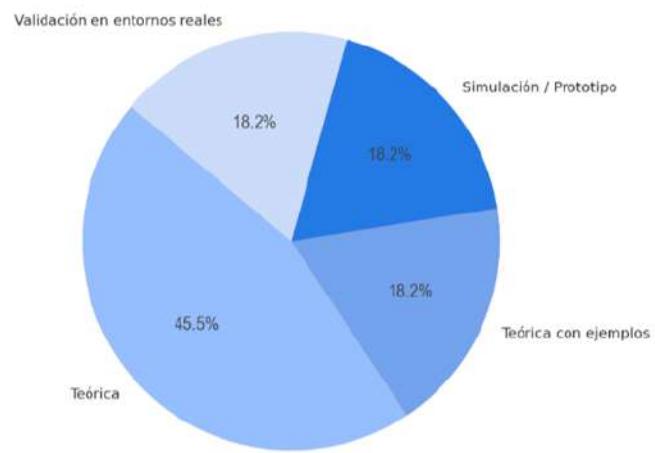
Aunque estas prácticas son heterogéneas en su formulación, convergen en la necesidad de incorporar controles proactivos desde las fases tempranas de diseño y no solo en la operación post-despliegue.

#### 3.1.4. Discusión Crítica

Uno de los hallazgos más relevantes es el hecho de que no existe ninguna estandarización entre los marcos de implementación, pues si bien algunos estudios se refieren a normas como ISO/IEC 27001, OWASP o NIST RMF (Calle-Méndez & Barriga-Andrade, 2025; Grechi et al., 2025), la mayoría de ellos proponen soluciones ad-hoc que son muy específicas a un contexto concreto. Además, únicamente cuatro de los estudios propuestos incluyen algún tipo de validación técnica (Alnami et al, 2025; Bai et al., 2025; Ahmed et al., 2025; Cajamarca et al., 2025), aunque centradas en IA o Cloud por separado, sugiriendo que las buenas prácticas no han sido suficientemente validadas. Esta falta de validación técnica será abordada en la fase experimental de la investigación a través del uso de un prototipo funcional. Además, casi la mitad de los estudios analizados cuenta con un enfoque meramente teórico (45,5%), mientras que solo una

minoría plantean soluciones mediante simulaciones, prototipos o entornos reales (cada uno con el 18,2%), lo cual es una muestra clara de la creciente necesidad de investigaciones que cuenten con validación empírica. Lo mencionado se observa a continuación en la Figura 6:

Frente a esto, algunos autores han planteado la posibilidad de una convergencia de marcos mediante arquitecturas híbridas que combinen, por ejemplo, el NIST AI RMF como marco de análisis de riesgos, controles técnicos del OWASP Top 10 for LLMs y complementando con los lineamientos de gestión de la ISO/IEC 27001 (Sharma & Dhiman, 2025). Así se combinaría



**Figura 6.** Nivel de validación de los estudios analizados.

Fuente: Los Autores

la gobernanza, el control técnico y la gestión de riesgos en un mismo enfoque, permitiendo una cobertura completa de proyectos de IA en la nube. Por otro lado, se observa la validez de enfoques complementarios como blockchain y cifrado avanzado que permiten cubrir las limitaciones clásicas en la protección de datos. Estudios como los de Lamar Peña et al. (2024), Vaca & Dulce-Villarreal (2024), Rueda-Castañeda et al. (2024), Dewangan y Chandrakar (2025) y Wazid et al. (2025) han demostrado la utilidad del blockchain para garantizar la trazabilidad, la autenticación segura y la emisión confiable de datos digitales, aportando valiosos mecanismos para fortalecer la integridad de modelos de IA desplegados en la nube. De manera similar, se han explorado mecanismos de autenticación biométrica multimodal en entornos cloud (Lin et al., 2023; Cui et al., 2025) y sistemas de detección de intrusiones basados en IA aplicados a infraestructuras críticas (Gujar, 2024), aunque estos enfoques siguen siendo específicos y carecen de validación conjunta en arquitecturas distribuidas de IA en la nube. Por último, se detecta un creciente interés por las preocupaciones éticas, pero aún no

están representadas en políticas técnicas claras, lo que representa una futura oportunidad de convergencia entre marcos normativos y principios éticos aplicables.

### 3.1.5. Propuesta Preliminar de Buenas Prácticas a validar

Como resultado de la revisión, se propone la siguiente lista preliminar de buenas prácticas para ser validadas en el prototipo experimental del estudio:

- Protección de API mediante tokens de acceso temporales.
- Autenticación multifactor y segregación de privilegios.
- Cifrado de datos sensibles en tránsito y en reposo.
- Registro detallado de eventos de seguridad para auditoría.

como el sistema diferencia entre usuario (user) y administrador (admin), permitiendo una gestión de privilegios efectiva. Además, el sistema de bloqueo automático de cuentas tras múltiples intentos fallidos de inicio de sesión y la trazabilidad de cada evento queda registrado en la base de datos, accesible mediante auditorías que podemos observar en la Figura 8. Por último, la Figura 9 detalla que de los 100 intentos con tokens JWT entre inválidos y expirados fueron 100% rechazados por parte del sistema, garantizando que las sesiones no se mantengan abiertas indefinidamente y previene ataques por fuerza bruta.

### 3.2.2 Registro de Eventos de Seguridad

El sistema realiza auditoría detallada de eventos críticos, incluyendo autenticaciones exitosas o fallidas, intentos de acceso desde IPs bloqueadas y uso de token inválidos o expirados. En las Figuras 10 y 11 se observan las validaciones de intentos

```

Status: 200 OK Size: 325 Bytes Time: 938 ms
Response Headers Cookies Results Docs
1 {
2   "access_token": "eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9
3     .eyJ1c2VyX2lkIjoxMSwidXNlcm5hbWUiOiJVc3VhcmlvMTAwIiwicm9sZSI6InVZX
4     IiLCJleHAiOjE3NTE1MzQwNTMsImhdCI6MTC1MTUwNTI1MywianRpIjoiSngzVE5Ca
5     mteUDnpjGRESDVoanRVSJ9
6   "token_type": "bearer",
7   "user": {
8     "username": "Usuario100",
9     "role": "user"
10   }
11 }
```

**Figura 7.** Validación en el inicio de sesión mediante Thunder Client (Credenciales Válidas) y generación del token JWT.

Fuente: Los Autores

| security_events |                  |         |            |  |         |        |                            |               |
|-----------------|------------------|---------|------------|--|---------|--------|----------------------------|---------------|
| id              | event_type       | user_id | ip_address | user_agent                                     | requ... | ris... | timestamp                  |               |
| 127             | SUCCESSFUL_LOGIN | 11      | 127.0.0.1  | Thunder Client (https://www.thunderclient.com) | NULL    | LOW    | 2025-07-03 01:58:40.979196 | Login exitoso |

**Figura 8.** Registro del evento de inicio exitoso en PostgreSQL.

Fuente: Los Autores

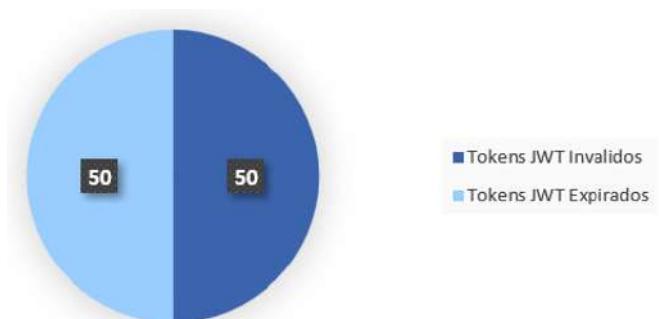
- Verificación de integridad mediante tecnologías distribuidas (ej. IOTA).
- Supervisión automática de tráfico mediante IA.
- Revisión de código y modelos bajo principios de transparencia.

Estas prácticas se incorporarán al prototipo con el fin de verificar su factibilidad técnica y analizar su efecto en la mitigación de riesgos dentro de un entorno de computación en la nube.

## 3.2 Resultados del Prototipo

### 3.2.1. Acceso al Sistema

En la Figura 7 se visualiza la validación del mecanismo de autenticación basado en tokens JWT con expiración automática, lo que asegura sesiones temporales y protegidas, también de



**Figura 9.** Distribución de 100 intentos fallidos de uso del chat con tokens JWT inválidos y expirados

Fuente: Los Autores



maliciosos y bloqueo de patrones de riesgos con respuestas automáticas. Todos los eventos son almacenados con su nivel de riesgo, usuario, IP y hora exacta. Al final, en la Figura 12 se muestra que, de 100 eventos aleatorios legítimos y maliciosos, fueron registrados en su totalidad con detalles proporcionando una trazabilidad completa para análisis de los datos y detención de amenazas.

rate\_limit (límite máximo 10 peticiones en un minuto por usuario) para mitigar abusos del sistema. En la Figura 17 se muestra la validación del método aplicado. Además, en la Figura 18 se demuestra que mediante pruebas de estrés y comparación de tiempos de respuesta promedio con respecto al límite de peticiones, se pudo obtener que dicho control contribuye significativamente a mantener un rendimiento óptimo del sistema, garantizando tiempos de respuesta estables.

The screenshot shows a POST request to <http://127.0.0.1:8000/chat>. The body contains JSON with a prompt field containing XSS code. The response status is 400 Bad Request, size is 64 Bytes, and time is 325 ms. The response body is: "detail": "Entrada rechazada por contener patrones sospechosos".

```

POST http://127.0.0.1:8000/chat
Status: 400 Bad Request Size: 64 Bytes Time: 325 ms
Response Headers Cookies Results Docs
1 {
2   "detail": "Entrada rechazada por contener patrones sospechosos"
3 }
  
```

**Figura 10.** Validación y clasificación de intentos maliciosos en el endpoint (/chat).

Fuente: Los Autores

The screenshot shows a POST request to <http://127.0.0.1:8000/chat>. The body contains JSON with a prompt field containing a shell command. The response status is 400 Bad Request, size is 58 Bytes, and time is 321 ms. The response body is: "detail": "Entrada bloqueada por políticas de seguridad".

```

POST http://127.0.0.1:8000/chat
Status: 400 Bad Request Size: 58 Bytes Time: 321 ms
Response Headers Cookies Results Docs
1 {
2   "detail": "Entrada bloqueada por políticas de seguridad"
3 }
  
```

**Figura 11.** Validación de detección y bloqueo de patrones de riesgo en el endpoint (/chat).

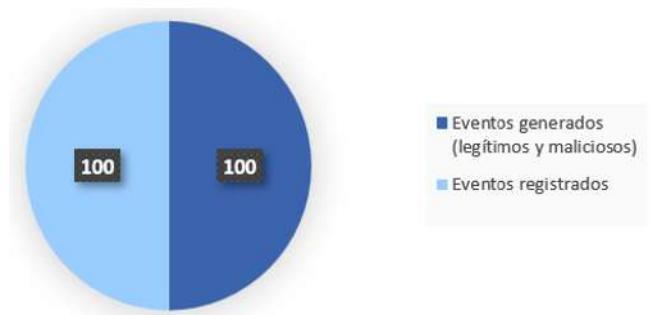
Fuente: Los Autores

### 3.2.3 Gestión de Vulnerabilidades

El sistema integra un módulo de detección de vulnerabilidades que analiza en tiempo real las solicitudes enviadas a la IA. Este módulo identifica patrones relacionados con inyecciones SQL, XSS, inyecciones de comandos y prompt injection en entradas del usuario como se muestra en la Figura 13. La solicitud puede ser bloqueada o permitida bajo advertencia dependiendo del nivel de riesgo detectado, pero siempre registrada en la base de datos. En la Figura 14 se detalla uno de los registros de Inyección SQL con Patrones maliciosos. Para terminar, la Figura 15 destaca que el sistema detectó y bloqueó el 100% de ataques aleatorios simulados clasificándolos por nivel de riesgo, así se minimiza la superficie de ataque mediante una respuesta automática, permitiendo la priorización de amenazas.

### 3.2.4. Controles de Red y Protección del Entorno

En la Figura 16 se detalla cómo se configuró el método check\_

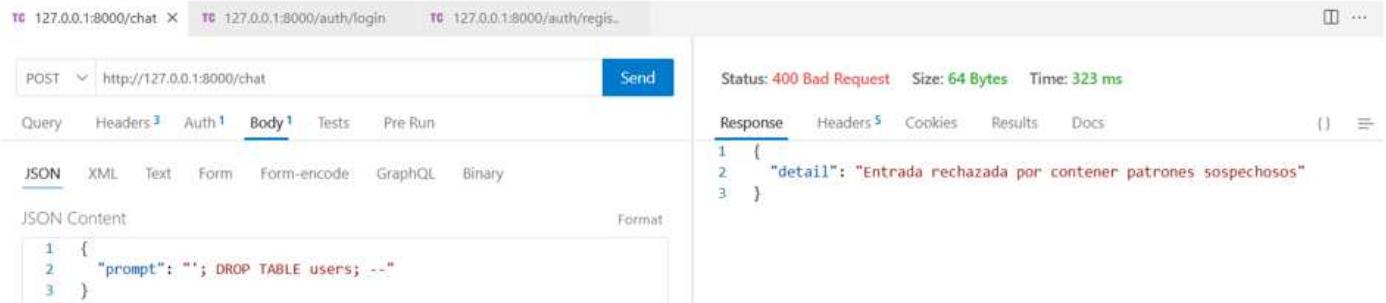


**Figura 12.** Generación de eventos vs el registro en la base de datos

Fuente: Los Autores

Se puede observar en la Figura 19 la lógica de bloqueo automático por nivel de riesgo y conjuntamente su validación en la Figura 20 con el bloqueo automático de dirección IP por exceder límite





```
POST http://127.0.0.1:8000/chat
Content-Type: application/json
{
  "prompt": "'"; DROP TABLE users; --"
}

Status: 400 Bad Request
Size: 64 Bytes
Time: 323 ms

Response Headers:
Content-Type: application/json
Content-Length: 64
Date: Mon, 10 Jul 2023 14:45:20 GMT
Connection: close
Server: Apache/2.4.41 (Ubuntu)
X-Powered-By: PHP/8.1.12-1+ubuntu22.04.1+deb.sury.org+1

{
  "detail": "Entrada rechazada por contener patrones sospechosos"
}
```

**Figura 13.** Validación y clasificación de intentos maliciosos en el endpoint (/chat).

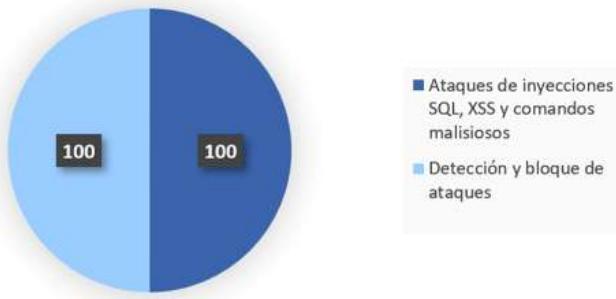
Fuente: Los Autores

138 CHAT\_RE... 13 127.0.0.1 Thunder Client (<https://www.thunderclient.co...> '; DROP TABLE users;... HIGH 2025-07... Vulnerabilidades: SQL\_INJECTION, DANGEROUS\_KEY

**Figura 14.** Registro de inyecciones SQL con patrones maliciosos en PostgreSQL.

Fuente: Los Autores

o generar entradas críticas. En la Figura 21 se puede visualizar que de 100 intentos aleatorios de entradas críticas el sistema bloqueó el 100% de las direcciones IP involucradas. Así mismo, se registró automáticamente en la base de datos, incluyendo detalles como la fecha, hora, dirección de IP del usuario y el contenido de la entrada maliciosa. Esto evidencia la efectividad del mecanismo de detección y respuesta ante amenazas.



**Figura 15.** Detección y bloqueo de inyecciones SQL, XSS y comandos maliciosos

Fuente: Los Autores

Los resultados confirman la hipótesis inicial del estudio: la implementación de buenas prácticas de seguridad si permiten fortalecer los sistemas de IA desplegados en la nube. En particular, se evidencia que estas prácticas establecen mecanismos efectivos de control de acceso, detección y mitigación de vulnerabilidades, auditoría de eventos y control del tráfico de red. Estas no solo incrementan la capacidad de los sistemas ante amenazas comunes, sino que también garantizan la integridad, disponibilidad y trazabilidad de los servicios de IA en la nube.

```
def check_rate_limit(self, ip_address: str, limit: int = 10, window: int = 60) -> bool:
    current_time = time.time()
    if ip_address not in self.rate_limit_storage:
        self.rate_limit_storage[ip_address] = []
    self.rate_limit_storage[ip_address].append(current_time)
    ts_for_ts in self.rate_limit_storage[ip_address]:
        if current_time - ts < window
    if len(self.rate_limit_storage[ip_address]) >= limit:
        return False
    self.rate_limit_storage[ip_address].append(current_time)
    return True
```

**Figura 16.** Método check\_rate\_limit.

Fuente: Los Autores

#### 4. Conclusiones

La integración de la inteligencia artificial en entornos de computación en la nube ha abierto nuevas oportunidades para la automatización, el procesamiento inteligente de datos y el desarrollo de soluciones escalables. No obstante, esta convergencia también ha traído amenazas importantes desde el punto de vista de la seguridad, en especial en la gestión de acceso a la información, la exposición de API y la protección de datos sensibles. El estudio, basado en una revisión sistemática de literatura científica reciente (2021-2025), ha determinado un conjunto representativo de buenas prácticas y mecanismos de protección que han sido propuestos para minimizar dichos riesgos, entre los que cabe mencionar: la autenticación fuerte, el cifrado, la auditoría de eventos y el control granular de los permisos. Como aporte práctico, se desarrolló un prototipo basado en inteligencia artificial, al que se le incorporaron algunas de las buenas prácticas identificadas. Este fue desplegado en un entorno Cloud, donde se aplicaron controles como tokens JWT, validación de entradas, cifrado en tránsito y reposo, y registro de eventos críticos. Si bien se trató de una simulación de alcance



The screenshot shows a Postman interface with three tabs at the top: '127.0.0.1:8000/chat' (Status: 200 OK), '127.0.0.1:8000/auth/login' (Status: 200 OK), and '127.0.0.1:8000/auth/regis...' (Status: 200 OK). The main area has a 'POST' method selected for 'http://127.0.0.1:8000/chat'. The 'Body' tab is active, showing JSON content: { "prompt": "Hola" }. The response panel shows a status of 429 Too Many Requests, size of 34 Bytes, and time of 320 ms. The response body is: { "detail": "Demasiadas peticiones" }.

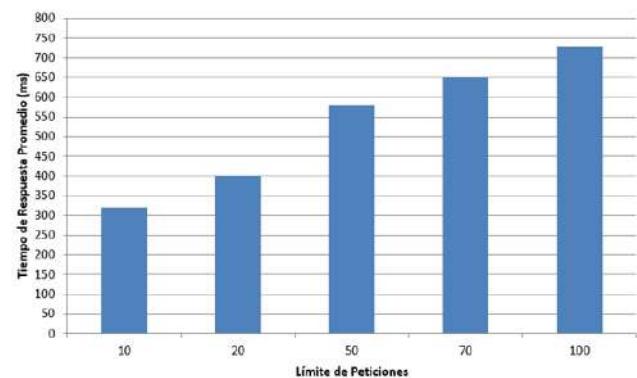
**Figura 17.** Validación del Método check\_rate\_limit en el endpoint (/chat).

Fuente: Los Autores

limitado, los resultados obtenidos permitieron comprobar que dichas prácticas son viables técnicamente y efectivas ante amenazas comunes en estos entornos, como inyecciones de comandos, abuso de endpoints o intentos de acceso masivo.

Este estudio pone de manifiesto que aplicar buenas prácticas de seguridad en proyectos de inteligencia artificial en la nube resulta, no sólo viable, sino necesario, incluso en las primeras etapas de prototipado. Además, refuerza la necesidad de contar con una guía fundamentada, basada en evidencia científica, que oriente a desarrolladores, ingenieros de seguridad y decisores en la puesta en práctica de los controles necesarios desde las primeras instancias de desarrollo. No obstante, se da considerando que existen limitaciones propias del uso de entornos simulados, por ello se sugiere que futuras investigaciones apunten a evaluar estos mecanismos en entornos de producción reales, donde juegan un papel relevante factores como la latencia, la escalabilidad, el rendimiento del sistema o la normativa vigente. También se recomienda incluir marcos éticos y de gobernanza para asegurar que se diseñen sistemas de inteligencia artificial seguros, pero también transparentes, auditables y que respeten la privacidad y

los derechos de los usuarios. Desde esta perspectiva, el presente artículo supone un primer paso en el diseño de sistemas de inteligencia artificial seguros en la nube, así como una base para futuras investigaciones que pretendan validar, extender o adaptar las buenas prácticas halladas en contextos más complejos, distribuidos y dinámicos.

**Figura 18.** Tiempo de respuesta con base al límite de peticiones.

```
if risk_level == "CRITICAL":
    security_manager.block_ip(client_ip)
    raise HTTPException(status_code=400, detail="Entrada bloqueada por políticas de seguridad")
```

**Figura 19.** Lógica de Bloqueo Automático por Nivel de Riesgo Crítico.

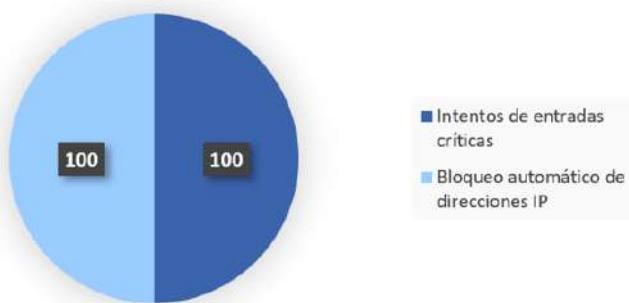
Fuente: Los Autores

The screenshot shows a Postman interface with three tabs at the top: '127.0.0.1:8000/chat' (Status: 200 OK), '127.0.0.1:8000/auth/login' (Status: 200 OK), and '127.0.0.1:8000/auth/regis...' (Status: 200 OK). The main area has a 'POST' method selected for 'http://127.0.0.1:8000/chat'. The 'Body' tab is active, showing JSON content: { "prompt": "ls; rm -rf / --force" }. The response panel shows a status of 400 Bad Request, size of 58 Bytes, and time of 324 ms. The response body is: { "detail": "Entrada bloqueada por políticas de seguridad" }.

**Figura 20.** Lógica de Bloqueo Automático por Nivel de Riesgo Crítico.

Fuente: Los Autores





**Figura 21.** Intentos de entradas críticas vs bloqueo automático de direcciones IP.

Fuente: Los Autores

### Contribución de los autores

**Yasbeck Jemima Mora Chávez:** Administración del proyecto, Investigación, Redacción y Metodología. **Reinaldo Benedicto Fernández González:** Investigación, y edición del artículo. **Joofre Antonio Honores Tapia:** Metodología, revisión, redacción y edición del artículo. **Milton Rafael Valarezo Pardo:** Metodología, revisión, redacción y edición del artículo.

### Declaración de conflictos de interés

Los autores manifiestan que no existe ningún conflicto de interés relacionado con este trabajo.

### Agradecimientos

Agradecemos a nuestros tutores por su acompañamiento académico y sus recomendaciones técnicas, que fueron decisivas en el desarrollo y mejora de esta investigación. También hacemos extensivas las gratitudes a nuestras familias y a las personas más significativas para nosotros, con el aliento que nos brindaron, lo cual tuvo un impacto sin duda considerable para que podamos cumplir con este proceso de investigación.

### Referencias bibliográficas

- Abioye, S. O., Oyedele, L. O., Akanbi, L., Ajayi, A., Davila Delgado, J. M., Bilal, M., Akinade, O. O., & Ahmed, A. (2021). Inteligencia artificial en el sector de la construcción: Una revisión de la situación actual, oportunidades y retos futuros. *Journal of Building Engineering*, 44, Article 103299. <https://doi.org/10.1016/j.jobe.2021.103299>
- Acosta Cortez, N. N. (2024). Impacto de la inteligencia artificial en la ciberseguridad empresarial: Un análisis crítico de la evolución de amenazas y medidas preventivas [Unpublished undergraduate thesis]. Universidad Técnica de Babahoyo, Facultad de Ciencias Informáticas. <https://dspace.utb.edu.ec/handle/49000/15738>
- Ahmed, W., Iqbal, W., Hassan, A., Ahmad, A., Ullah, F., & Srivastava, G. (2025). Elevating e-health excellence with IOTA distributed ledger technology: Sustaining data integrity in next-gen fog-driven systems. *Future Generation Computer Systems*, 168, Article 107755. <https://doi.org/10.1016/j.future.2024.107755>
- Alnami, H., Mahgoub, I., Al-Najada, H., & Alalwany, E. (2025). A distributed machine learning-based scheme for real-time highway traffic flow prediction in Internet of Vehicles. *Future Internet*, 17(3), Article 131. <https://doi.org/10.3390/fi17030131>
- Bai, Y., Zhao, H., Shi, X., & Chen, L. (2025). Towards practical and privacy-preserving CNN inference service for cloud-based medical imaging analysis: A homomorphic encryption-based approach. *Computer Methods and Programs in Biomedicine*, 261, Article 108599. <https://doi.org/10.1016/j.cmpb.2024.108599>
- Bello, S. A., Oyedele, L. O., Akinade, O. O., Bilal, M., Davila Delgado, J. M., Akanbi, L. A., Ajayi, A. O., & Owolabi, H. A. (2021). Cloud computing in construction industry: Use cases, benefits and challenges. *Automation in Construction*, 122, Article 103441. <https://doi.org/10.1016/j.autcon.2020.103441>
- Cajamarca Sacoto, J. C., Chuquin Machangara, O. M., & Recalde Araujo, J. C. (2025). El papel de la inteligencia artificial en la ciberseguridad de redes empresariales. *Revista Retos Para La Investigación*, 4(1), 65–82. <https://doi.org/10.62465/rri.v4n1.2025.126>
- Calle-Méndez, J. L., & Barriga-Andrade, J. J. (2025). Amenazas de seguridad asociadas con la integración de inteligencia artificial en sistemas de información: *Revisión sistemática*. *MQRInvestigar*, 9(1), Article e128. <https://doi.org/10.56048/MQR20225.9.1.2025.e128>
- Chen, H.-P., & Ying, K.-C. (2022). Inteligencia artificial en el sector de la construcción: Principales trayectorias de desarrollo y perspectivas de futuro. *Applied Sciences*, 12(12), Article 5832. <https://doi.org/10.3390/app12125832>
- Cui, W., Lin, Q., Shi, J., Zhou, X., Li, Z., Zhan, H., & Zhang, L. (2025). A secure and efficient framework for multimodal biometric authentication in cloud computing. *Applied*

- Sciences, 15(7), Article 3827. <https://doi.org/10.3390/app15073827>
- Dewangan, N. K., & Chandrakar, P. (2025). TreatChain: A patient-centric treatment cycle security system using blockchain and AI in the cloud. *Peer-to-Peer Networking and Applications*, 18(3), Article 154. <https://doi.org/10.1007/s12083-025-01984-z>
- Flores-Cedeño, P. R., Zambrano-Pilay, E. C., & Chiriboga-Mendoza, F. R. (2024). Seguridad informática e inteligencia artificial en la investigación científica. *Revista Científica INGENIAR: Ingeniería, Tecnología e Investigación*, 7(13), 2–10. <https://www.journalingeniar.org/index.php/ingeniar/article/view/177>
- Grechi, V. L., de Oliveira, A. L., & Braga, R. T. V. (2025). Model-driven safety and security co-analysis: A systematic literature review. *Journal of Systems and Software*, 220, Article 112251. <https://doi.org/10.1016/j.jss.2024.112251>
- Gujar, S. S. (2024, December). AI-enhanced intrusion detection systems for strengthening critical infrastructure security [Conference session]. 2024 Global Conference on Communications and Information Technologies (GCCIT), Bangalore, India. <https://doi.org/10.1109/GCCIT63234.2024.10861950>
- Gutiérrez Rodríguez, J. D., & Castellanos-Sánchez, M. (2023). Transparencia algorítmica y estado abierto en Colombia. *Reflexión Política*, 25(52), 6–21. <https://doi.org/10.29375/01240781.4789>
- Jiang, Y., Ye, Y., Zhao, H., Zhang, S., Cao, Y., & Gu, J. (2021). Analysis of smart water conservancy. *Shuili Xuebao/Journal of Hydraulic Engineering*, 52(11), 1355–1368. <https://doi.org/10.13243/j.cnki.slxb.20210633>
- Kyivska, K. I., & Tsutsiura, S. (2021). Implementation of artificial intelligence in the construction industry and analysis of existing technologies. *Technology Audit and Production Reserves*, 2(2), Article 58. <https://doi.org/10.15587/2706-5448.2021.229532>
- Lamar Peña, F. S., Vega Mite, G. A., Honores Tapia, J. A., & Cárdenas Villavicencio, O. E. (2024). Validación y emisión de certificados en educación superior utilizando tecnología blockchain. *Informática y Sistemas: Revista de Tecnologías de la Informática y las Comunicaciones*,
- 8(1), 36–51. <https://doi.org/10.33936/isrtic.v8i1.6535>
- Lin, L., Chen, C., Pan, L., Zhang, L. Y., Wang, Z., Zhang, J., & Xiang, Y. (2023). A survey of PPG's application in authentication. *Computers & Security*, 135, Article 103488. <https://doi.org/10.1016/j.cose.2023.103488>
- Pan, Y., & Zhang, L. (2021). Roles of artificial intelligence in construction engineering and management: A critical review and future trends. *Automation in Construction*, 122, Article 103517. <https://doi.org/10.1016/j.autcon.2020.103517>
- Rueda-Castañeda, J. E., Gallego-Gómez, N., Estanling-Cárdenes, E., Tello, J. S., & García-Pineda, V. (2024). Identificación de variables relacionadas a la seguridad informática desde la aplicación de la tecnología blockchain. *Revista Politécnica*, 20(40), 9–29. <https://doi.org/10.33571/rpolitec.v20n40a1>
- Sharma, N., & Dhiman, P. (2025). A survey on IoT security: Challenges and their mitigation strategies in cloud-centric architectures. *Cluster Computing*. Advance online publication. <https://doi.org/10.1007/s10586-025-05208-0>
- Vaca, P. A., & Dulce-Villarreal, E. R. (2024). Blockchain para asegurar la integridad y trazabilidad en la cadena de custodia de evidencia digital en informática forense: *Un estudio de mapeo sistemático*. *TecnoLógicas*, 27(60), Article e3049. <https://doi.org/10.22430/22565337.3049>
- Wazid, M., Mittal, S., Das, A. K., Islam, S. H., & Alenezi, M. (2025). Designing secure blockchain-based authentication and authorization mechanisms for smart city e-health systems. *Journal of Systems Architecture*, 156, Article 103365. <https://doi.org/10.1016/j.sysarc.2025.103365>
- Zhang, Q., Cheng, L., & Boutaba, R. (2010). Computación en la nube: Estado del arte y desafíos de la investigación. *Journal of Internet Services and Applications*, 1(1), 7–18. <https://doi.org/10.1007/s13174-010-0007-6>
- Zhang, X., Antwi-Afari, M. F., Zhang, Y., & Xing, X. (2024). The impact of artificial intelligence on organizational justice and project performance: A systematic literature and science mapping review. *Buildings*, 14(1), Article 259. <https://doi.org/10.3390/buildings14010259>





## Balanceo de Conjuntos de Datos Basado en Redes Generativas Aplicado a Imágenes del Sector Agrícola

## Generative Network-Based Data Balancing Applied to Agricultural Image Datasets

### Autores

\*Luis Jesús Montesdeoca Espinoza 

 luis.montesdeoca.41@espm.edu.ec

Stalin Joel Zambrano Rojas 

 stalin.zambrano@espm.edu.ec

Victor Joel Pinargote Bravo 

 vpinargote@espm.edu.ec

Luis Cristobal Cedeño Valarezo 

 lcedeno@espm.edu.ec

Escuela Superior Politécnica  
Agropecuaria de Manabí Manuel Félix  
López, Calceta, Manabí, Ecuador.

\*Autor para correspondencia

### Comó citar el artículo:

Montesdeoca Espinoza, L.J., Zambrano Rojas, S.J., Pinargote Bravo, V.J. & Cedeño Valarezo, L.C. (2025). Balanceo de Conjuntos de Datos Basado en Redes Generativas Aplicado a Imágenes del Sector Agrícola. *Informática y Sistemas* 9(2), pp. 164-176. <https://doi.org/10.33936/isys.v9i2.7782>

Enviado: 11/08/2025

Aceptado: 12/09/2025

Publicado: 12/09/2025

### Resumen

El desbalance de clases en conjuntos de datos agrícolas limita el rendimiento de los modelos de clasificación basados en redes neuronales convolucionales, al dificultar la identificación precisa de clases minoritarias. Con el objetivo de mitigar este problema, se adaptó la metodología CRISP-DM para generar datos sintéticos mediante redes generativas adversarias Wasserstein (WGAN-GP), utilizando defectos segmentados de enfermedades del aguacate (Scab y Anthracnose) extraídos con técnicas de visión por computador. Estas anomalías fueron integradas en imágenes de frutos sanos para construir un conjunto de datos equilibrado. Posteriormente, se entrenó un modelo de clasificación utilizando la arquitectura InceptionV3 con aprendizaje por transferencia, evaluando su desempeño tanto con el conjunto desequilibrado como con el conjunto balanceado. Los resultados evidenciaron mejoras notables en la precisión, especialmente en la clasificación de enfermedades, cuando se utilizó el conjunto equilibrado, alcanzando una precisión de validación del 97.74%. Este estudio demuestra que el uso de datos sintéticos puede ser una solución efectiva para mejorar la capacidad predictiva de modelos en contextos donde la recopilación de datos reales es limitada o costosa.

**Palabras clave:** Aprendizaje profundo; Balanceo de clases; Datos sintéticos; Redes convolucionales; Segmentación de imágenes

### Abstract

The imbalance of classes in agricultural datasets limits the performance of classification models based on convolutional neural networks, as it makes it difficult to accurately identify minority classes. To mitigate this problem, the CRISP-DM methodology was adapted to generate synthetic data using Wasserstein generative adversarial networks (WGAN-GP), using segmented defects from avocado diseases (Scab and Anthracnose) extracted with computer vision techniques. These anomalies were integrated into images of healthy fruits to construct a balanced dataset. The InceptionV3 architecture with transfer learning was then used to train a classification model, and its performance was assessed using both the balanced and unbalanced datasets. Using the balanced dataset resulted in significant accuracy gains, particularly in disease classification, with a validation accuracy of 97.74%. This study shows that in situations where real data collection is expensive or limited, using synthetic data can be a useful way to increase the predictive power of models.

**Keywords:** Deep learning; Class balancing; Convolutional networks; Synthetic data; Image segmentation



## 1. Introducción

La agricultura es un pilar esencial para la seguridad alimentaria y el desarrollo económico global (Thayer et al., 2020). Sin embargo, enfrenta desafíos crecientes como el cambio climático, la degradación del suelo y la propagación de enfermedades en cultivos, que demandan soluciones tecnológicas innovadoras. La Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, 2021) ha resaltado la necesidad de integrar herramientas digitales para lograr sistemas agrícolas más resilientes y eficientes.

El aprendizaje automático, una de las tecnologías emergentes, ha demostrado tener un alto potencial para optimizar los procesos productivos, ayudar a la toma de decisiones fundamentadas en datos y mejorar la detección oportuna de enfermedades (Soria Olivas et al., 2023; Shalev-Shwartz & Ben-David, 2014). En el tratamiento de imágenes agrícolas, se han empleado en gran medida las arquitecturas codificador-decodificador, U-Net y las redes neuronales convolucionales (CNN), con resultados sobresalientes en la clasificación y segmentación de hojas y frutos (Lei et al., 2024). Sin embargo, su desempeño tiende a estar limitado por la disponibilidad de datos equilibrados y representativos.

Investigaciones anteriores han indagado sobre la aplicación de las redes generativas adversarias (GAN) para incrementar las clases minoritarias en fotografías agrícolas, alcanzando avances en sensibilidad y precisión (Lu et al., 2022; Carvajal Chávez, 2023). No obstante, estos estudios tienen limitaciones típicas: el riesgo de sobreajuste al volver a utilizar imágenes de validación en el proceso de generación, la escasa evaluación de la variabilidad morfológica que se produce y la falta de control sobre cómo se integran los datos sintéticos. Asimismo, la mayor parte se ha enfocado en enfermedades o cultivos concretos sin especificar la posibilidad de reproducir sus métodos.

En los años recientes (2023-2025), la literatura sobre visión por computadora en el sector agrícola ha investigado tácticas de rebalanceo de clases que combinan técnicas de oversampling con datos sintéticos. En concreto, numerosos estudios indican que los modelos generativos (GANs y sus variantes) tienen la capacidad de incrementar la diversidad morfológica de las lesiones y optimizar la identificación de clases minoritarias, lo cual tiene efectos positivos en el recall/precisión y F1 en cultivos y enfermedades foliares ampliamente investigadas como el tomate (Wang et al., 2025).

Este conjunto de pruebas más recientes en el ámbito agrícola amplía y complementa las revisiones anteriores acerca de GANs en la agroalimentación, destacando la importancia del empleo de métricas de calidad de síntesis (como FID e IS) para verificar la utilidad que tienen los datos producidos en tareas posteriores a la clasificación (Wang et al., 2025).

A diferencia de estudios previos que aplican GANs de manera genérica para aumentar imágenes agrícolas, el aporte central de este trabajo consiste en entrenar una WGAN-GP

exclusivamente con defectos segmentados de enfermedades reales de aguacate (Scab y Anthracnose) y posteriormente insertarlos de forma controlada en frutos sanos. Esta estrategia garantiza que las imágenes sintéticas conserven la textura, iluminación y fondo natural del fruto, al tiempo que diversifican la morfología de las lesiones. Con ello se logra un balanceo más realista y contextualizado del conjunto de datos, evitando las duplicaciones artificiales propias de técnicas clásicas como oversampling o SMOTE. Además, al definir la partición de los datos antes de la generación, se controla la fuga de información entre conjuntos y se asegura la reproducibilidad del pipeline. Este enfoque, complementado con la adaptación explícita de la metodología CRISP-DM y la evaluación con métricas robustas por clase, constituye la principal innovación metodológica frente a propuestas anteriores en el área.

Se propone como hipótesis que la utilización de datos sintéticos producidos por WGAN-GP para el balanceo de clases podría incrementar la habilidad de generalización de modelos que clasifican imágenes agrícolas, en lugar del entrenamiento con conjuntos desbalanceados. El propósito general es crear y analizar un flujo de trabajo que pueda repetirse, incluyendo la extracción de defectos, la producción controlada de imágenes sintéticas y la evaluación del impacto que tiene en el rendimiento de un modelo basado en InceptionV3.

## 2. Materiales y Métodos

La investigación se desarrolló entre abril y julio de 2025 en la Unidad de Docencia Investigación y Vinculación (UDIV), perteneciente a la Escuela Superior Politécnica Agropecuaria de Manabí “Manuel Félix López” (ESPAM MFL), Ecuador. El laboratorio cuenta con estaciones de trabajo equipadas con GPU NVIDIA RTX 3060 (12 GB) y entornos configurados con Python 3.10, TensorFlow 2.x, OpenCV y JupyterLab. El tiempo total de cómputo fue de aproximadamente 120 horas, distribuidas entre preparación de datos, segmentación de defectos, entrenamiento de redes generativas adversarias (WGAN-GP) y entrenamiento de modelos clasificadores mediante aprendizaje por transferencia.

### 2.1 Conjunto de Datos y Caracterización

El conjunto de datos utilizado corresponde a imágenes de aguacates capturadas en un estudio fotográfico controlado en Morelos, México, durante un único ciclo productivo, disponible en el repositorio de Kaggle llamado “Clasificación de enfermedades del aguacatero” del usuario camposfe1, este es un repositorio abierto a todo público con autorización para que se haga el uso adecuado de esa colección de imágenes. Las imágenes están clasificadas en tres condiciones sanitarias: Healthy (sano), Scab y Anthracnose, con un total de 3983 registros distribuidos en

1764, 1197 y 1022 imágenes respectivamente. Esta distribución presenta un desbalance moderado (proporciones aproximadas del 44%, 30% y 26%), que podría afectar el rendimiento de los modelos de clasificación, particularmente en la identificación de clases minoritarias. También en este mismo repositorio existe una carpeta test, esta carpeta no está clasificada en clases. Esta carpeta se utilizará para comprobar la efectividad del modelo con 100 imágenes (estos datos no se usarán para el entrenamiento será exclusivamente para la validación del modelo).

En el repositorio menciona que las condiciones de captura fueron estandarizadas para minimizar la variabilidad en iluminación, enfoque y encuadre. La caja de estudio utilizada tiene dimensiones de  $22 \times 24 \times 23$  cm, con una intensidad lumínica de 450 lm, fondo blanco y colocación central del fruto.

Así mismo también menciona que las imágenes se capturaron con cámara digital en formato jpg con las siguientes especificaciones: resolución de  $4496 \times 3000$  pixeles, relación de aspecto 4:3, apertura f/3.8, velocidad de obturación 1/100 s, valor ISO 200, distancia focal 24 mm, compensación de exposición 0 paso y sin uso de flash. La iluminación estable y el fondo uniforme permitieron obtener imágenes de alta calidad.

Para evitar fuga de información hacia el conjunto de validación, la generación de imágenes sintéticas se realizó únicamente a partir de los ejemplos incluidos en el conjunto de entrenamiento, definiendo las particiones antes del entrenamiento de la WGAN-GP. Las imágenes están organizadas según el archivo labels.csv como se muestra en la Tabla 1, que estructura el conjunto de datos en formato multiclase mediante los siguientes atributos: *Identification* (identificador único de cada imagen), *Condition* (etiqueta de clase principal), y tres columnas binarias *Healthy*, *Scab*, *Anthracnose* que indican la pertenencia a cada clase.

**Tabla 1.** Tabla representativa del archivo labels.csv.

Fuente: Los Autores.

| Identification | Condition | Healthy | Scab | Anthracnose |
|----------------|-----------|---------|------|-------------|
| Avocado 001    | Scab      | 0       | 1    | 0           |
| Avocado 002    | Scab      | 0       | 1    | 0           |
| Avocado 003    | Healthy   | 1       | 0    | 0           |
| Avocado 004    | Scab      | 0       | 1    | 0           |

La enfermedad Scab es causada por el hongo *Sphaceloma perseae* y se manifiesta mediante lesiones oscuras y corchosas en la superficie del fruto, mientras que la Anthracnose, provocada por *Colletotrichum gloeosporioides*, se presenta con manchas oscuras que evolucionan a lesiones necróticas profundas, afectando significativamente la calidad del fruto (Bustamante et al., 2022).

## 2.2 Segmentación con OpenCV

Se implementaron técnicas de segmentación con la biblioteca OpenCV para entrenar una red generativa con el fin de simular efectos visuales realistas de patologías. Para obtener las zonas de las enfermedades presentes en imágenes de aguacate con Scab y Anthracnose, se usaron máscaras morfológicas, algoritmos de umbralización y transformaciones cromáticas dentro de los espacios Lab y HSV.

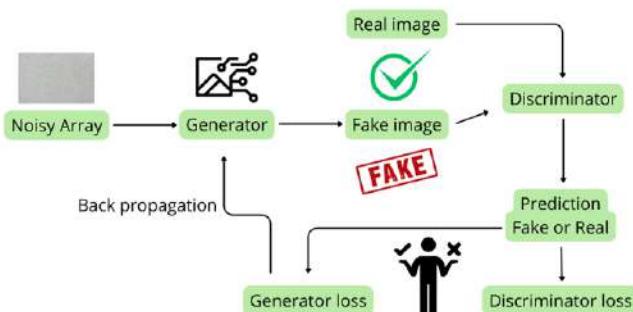
Se ha comprobado que estas técnicas son particularmente eficaces para dividir lesiones foliares y zonas enfermas a través de umbralización utilizando el canal H del espacio HSV, lo cual mejora la precisión sin necesidad de modelos pesados (Hai et al., 2025). De igual forma, herramientas de procesamiento parecidas se han empleado con éxito para automatizar la cuantificación de los síntomas de patologías y luchar contra las fluctuaciones en la iluminación a través de segmentación binaria adaptativa (Paauw et al., 2024).

## 2.3 Arquitectura del Modelo Generativo (WGAN-GP)

La WGAN-GP no realiza una tarea supervisada ni se evalúa con métricas de clasificación; su finalidad es aprender la distribución morfológica de los defectos segmentados (máscaras binarias de Scab/Anthracnose) y sintetizar nuevas instancias de defecto. Por esta razón, no se definió un particionado train/val/test para la GAN: su objetivo no es “generalizar” a datos no vistos, sino producir parches de lesión plausibles que luego se insertan sobre frutos sanos.

La arquitectura de la red generativa utilizada fue una WGAN como se muestra en la Figura 1 con penalización de gradiente (WGAN-GP), compuesta por:

- **Generador:** red secuencial con capas Dense, BatchNormalization, LeakyReLU y Conv2DTranspose, finalizando con activación tanh.
- **Discriminador (Crítico):** arquitectura espejo inversa con capas Conv2D, LeakyReLU y Dropout, sin activación final (por propiedad de la función de pérdida Wasserstein).



**Figura 1.** Diagrama de la arquitectura WGAN-GP mostrando el flujo de datos y retroalimentación.

Fuente: Los Autores.



Se emplearon las máscaras obtenidas con OpenCV como fundamento para entrenar una red generativa adversarial de Wasserstein (WGAN) que se enfoca únicamente en la creación de imperfecciones (necrosis, manchas, escoriaciones). La red, después de ser entrenada, posibilitó la síntesis de nuevos defectos, que se superpusieron posteriormente en imágenes aleatorias de aguacates sanos. De este modo, se generó una representación sintética de frutas enfermas. Esta estrategia posibilitó que la representación de las clases minoritarias crezca de forma controlada, sin modificar la distribución de la clase saludable (Salem et al., 2020).

### HiperParámetros de la WGAN-GP

Para garantizar la reproducibilidad del modelo generativo, en la Tabla 2 se resumen los principales hiper parámetros utilizados en el entrenamiento de la WGAN-GP. Estos valores fueron seleccionados siguiendo las configuraciones recomendadas en la literatura y ajustados de forma preliminar mediante pruebas exploratorias.

**Tabla 2.** HiperParámetros utilizados en la WGAN-GP.

Fuente: Los Autores.

| Parámetro   | Valor   |
|---|---|
| Dimensión del vector de ruido (z)                             | 100   |
| Optimizador   | Adam ( $\beta_1=0.5$ , $\beta_2=0.9$ )  |
| Tasa de Aprendizaje   | 0.0001  |
| Batch size  | 32  |
| Épocas  | 5000  |
| Penalización de gradiente ( $\lambda$ )                       | 10  |
| Número de actualizaciones del crítico ( $n_{\text{critic}}$ ) | 5   |
| Función de activación (generador)                             | tanh  |
| Función de activación (crítico)                               | LeakyReLU (sin capa final)<br>Dense( $4 \times 4 \times 256$ ) → BN → LeakyReLU → Conv2DTrans(128, $k=4$ , $s=2$ ) → BN → LeakyReLU → Conv2DTrans(64, $k=4$ , $s=2$ ) → BN → LeakyReLU → Conv2DTrans(3, $k=3$ , $s=1$ , act=tanh) (ajusta canales/tamaño real)<br>Conv2D(64, $k=4$ , $s=2$ ) → LeakyReLU → Conv2D(128, $k=4$ , $s=2$ ) → LeakyReLU → Conv2D(256, $k=4$ , $s=2$ ) → Flatten → Dense(1) |
| Generador   |   |
| Critic  |   |
| Selección de checkpoint                                       | Se guardó cada 200 épocas;<br>y se eligió el mejor modelo manualmente (el modelo que fue más regular generando los defectos sintéticos).  |

### Política de Partición de Datos

En el caso particular del entrenamiento de la WGAN-GP, no se implementó una política de partición tradicional en conjuntos de entrenamiento, validación y prueba, como suele hacerse en tareas de aprendizaje supervisado. El objetivo de este modelo generativo no es medir su rendimiento en datos no vistos, sino aprender la distribución estadística de los defectos segmentados de aguacate para poder sintetizar nuevas instancias realistas. En consecuencia, todas las imágenes disponibles de defectos fueron utilizadas para el entrenamiento, aplicando previamente técnicas de aumento de datos (rotaciones, variación de brillo y espejado).

De esta manera, la “fuga de información” entendida como el uso indebido de datos de validación o prueba en el entrenamiento no representa un riesgo en esta etapa, dado que la WGAN-GP no se evalúa mediante métricas de clasificación convencionales. La validación del enfoque se llevó a cabo posteriormente, en la fase de clasificación con Inception V3, donde sí se aplicó una división explícita de los datos en entrenamiento, validación y prueba para garantizar la independencia entre los conjuntos y asegurar la reproducibilidad del experimento.

A su vez se logró controlar qué frutos del mismo origen no aparezcan en conjuntos distintos realizando una partición por especie controlando los datos mediante carpetas separadas.

### 2.4 Clasificación con Red Convolucional InceptionV3

#### Arquitectura del modelo

Para analizar el efecto de equilibrar clases a través de imágenes sintéticas generadas, se utilizó una red neuronal convolucional con la metodología de transfer learning, que tiene como fundamento la arquitectura InceptionV3 y fue entrenada previamente con el conjunto de datos ImageNet. Esta arquitectura profunda ha demostrado ser muy eficaz en la clasificación de imágenes naturales, lo que posibilita aprovechar el conocimiento previamente adquirido y agilizar el proceso de entrenamiento, además de optimizar el rendimiento (Hossen et al., 2025).

Para prevenir el sobreajuste, se implementó la técnica EarlyStopping con patience de 5 y restore\_best\_weights=True en la configuración. Esto posibilitó parar el entrenamiento antes de tiempo. En la situación con datos desbalanceados, esta estrategia fue crucial porque el entrenamiento se detuvo sin intervención humana en la época 24. Esto indica que el modelo llegó a su mejor rendimiento antes de finalizar las 50 épocas programadas, lo cual optimizó el tiempo de cómputo y mantuvo la capacidad de generalización.

La arquitectura del modelo base de InceptionV3 se muestra en la Figura 2.

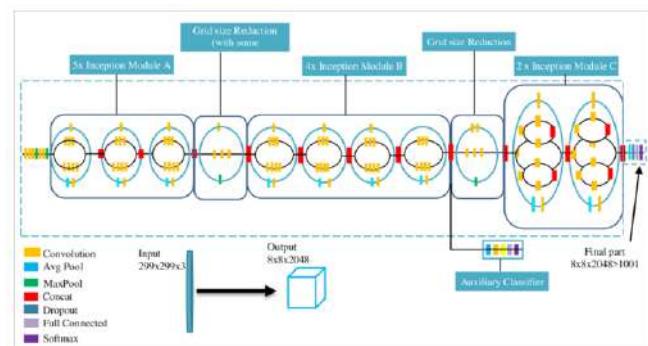
#### Fase 1 – Transfer Learning:

- Modelo base: InceptionV3 preentrenado en ImageNet.
- Congelación de capas: todas las capas convolucionales del modelo base fueron congeladas.
- Optimizador: Adam (tasa de aprendizaje por defecto).
- Función de pérdida: *categorical\_crossentropy*.
- *EarlyStopping*: habilitado con *patience=5* y *restore\_best\_weights=True*.

#### Fase 2 – Fine-Tuning:

- Se descongelaron las últimas 50 capas del modelo base.
- Épocas: 50 (Programadas).
- *EarlyStopping* mantenido.

Una capa final de salida *Dense* (3, *activation='softmax'*) para realizar la clasificación multiclase en tres categorías: *Healthy*, *Scab* y *Anthracnose*.



**Figura 2.** Estructura completa de la red InceptionV3 con sus bloques de inception y clasificador auxiliar.

Fuente: (Iparraguirre-Villanueva et al., 2022)

#### Configuración de entrenamiento

El modelo final fue compilado con los siguientes parámetros:

- Optimizador: Adam.
- Función de pérdida: *categorical\_crossentropy*, adecuada para clasificación multiclase.
- Métrica: Precisión (*accuracy*).

El entrenamiento se realizó durante 50 épocas programadas

con un tamaño de *batch* de 32, utilizando imágenes de tamaño 224x224 píxeles y aceleración mediante GPU.

#### Preprocesamiento de los datos

Se utilizó la clase *ImageDataGenerator* para normalizar las imágenes (*rescale=1./255*) y dividir automáticamente el conjunto de datos en:

- Entrenamiento: 80% de las imágenes.
- Validación: 20% de las imágenes.

Las imágenes utilizadas provenían de la carpeta datos-balanceados y datos-desbalanceados, la cuales contienen tres subdirectorios representando las clases correspondientes. Se observó el comportamiento del modelo y, durante el entrenamiento, se visualizaron las curvas de pérdida y precisión para las dos fases (entrenamiento y validación), lo cual posibilitó identificar potenciales problemas como el sobreajuste.

Finalmente, los modelos entrenados fueron almacenados para futuras evaluaciones o integraciones en sistemas de diagnóstico automatizado.

La división en entrenamiento (80%) y validación (20%) se realizó mediante el parámetro *validation\_split* de la clase *ImageDataGenerator* en Keras/TensorFlow, lo cual implica que la separación fue aleatoria a nivel global y no estratificada por clase. En consecuencia, las proporciones exactas de cada clase en entrenamiento y validación pueden variar ligeramente respecto a la distribución original del conjunto. Esta estrategia fue seleccionada debido a que:

- Permite mantener un flujo automatizado y reproducible en Keras sin necesidad de particiones manuales.
- Evita posibles sesgos introducidos por una estratificación rígida, dado que el objetivo principal del estudio era evaluar el impacto global del balanceo con imágenes sintéticas.
- Se complementa con un conjunto de prueba completamente independiente (100 imágenes de la carpeta test del repositorio Kaggle), lo que asegura la evaluación fuera del dominio de entrenamiento y validación.

#### 2.5 Diseño Experimental y Comparación

Para evaluar el impacto del balanceo conjunto de datos mediante imágenes sintéticas, se diseñó un experimento comparativo empleando una red neuronal convolucional con arquitectura InceptionV3, entrenada bajo dos condiciones distintas:

- Conjunto de datos original desbalanceado.
- Conjunto de datos balanceado mediante imágenes generadas con WGAN.



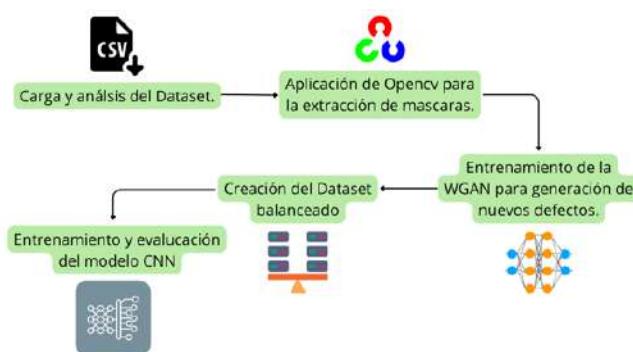
Ambos modelos fueron entrenados utilizando los mismos parámetros de arquitectura, optimización y preprocesamiento, lo que garantizó condiciones experimentales equivalentes y permitió una comparación objetiva del desempeño.

### Métrica de evaluación

En esta investigación, la métrica principal utilizada para evaluar la eficiencia del modelo de clasificación fue la precisión (accuracy), dado que el objetivo principal consistía en comparar el rendimiento entre un conjunto de datos desbalanceado y otro balanceado con imágenes sintéticas. Adicionalmente, para reforzar la validez del enfoque y confirmar la capacidad de generalización, se consideró la evaluación en un conjunto de prueba independiente no visto durante el entrenamiento ni la validación, reportando métricas de desempeño por clase (precisión, recall y F1-score).

### 2.5 Arquitectura General del Flujo Experimental

El flujo general del sistema se estructuró en las siguientes etapas que se puede ver en la Figura 3. Este diseño garantiza la replicabilidad del experimento y puede adaptarse a otros cultivos o enfermedades agrícolas.



**Figura 3.** Flujo de trabajo completo del sistema: desde carga de datos hasta entrenamiento de modelos.

Fuente: Los Autores.

El código, las configuraciones y los archivos de división (CSV) se encuentran disponibles en <https://github.com/JesusMe001/Balanceo-de-Conjuntos-de-Datos-Basados-en-Redes-Generativas-Aplicado-a-Im genes-del-Sector-Agrcola.git>, incluyendo ejemplos representativos de máscaras y sintéticos. Se documentan versiones de librerías y modelos ya generados listos para usarse.

### 3. Resultados y Discusión

#### 3.1 Preparación del Conjunto de Datos y Segmentación de Defectos

El procedimiento inicial consistió en clasificar el conjunto de datos original de imágenes de aguacates en tres carpetas, cada una de las cuales correspondía a sus respectivas clases: Scab, Anthracnose y Healthy. Esta estructura facilita un flujo automatizado para cargar y procesar imágenes, ya sea para la clasificación o para las labores de segmentación.

Para aislar solamente las zonas afectadas por enfermedades en las imágenes que se han marcado como Scab y Anthracnose, se llevó a cabo un proceso personalizado de extracción de defectos, empleando la biblioteca OpenCV. Este procedimiento fue esencial para crear un conjunto de datos realista y enfocado, el cual se usará como entrada para entrenar el modelo generativo WGAN.

El procedimiento incluyó los siguientes pasos:

**Conversión a espacio de color HSV:** para facilitar la segmentación según tonalidades específicas del fruto y su fondo.

**Definición de máscaras de exclusión:**

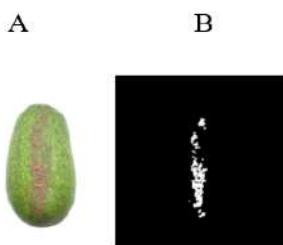
- Se creó una máscara para eliminar el fondo blanco (valores altos en el canal V).
- Se definió otra máscara para excluir el verde sano del aguacate (tonos entre 35° y 90° en el canal H).

**Combinación de máscaras:** se unieron ambas zonas a excluir y se invirtió la máscara, conservando únicamente los elementos que no son ni blancos ni verdes, es decir, las áreas enfermas.

**Limpieza morfológica:** se aplicó una operación de apertura para reducir ruido y mejorar la definición del área segmentada.

**Extracción final del defecto:** se aplicó la máscara obtenida sobre la imagen original, obteniendo una representación aislada del defecto, lista para ser utilizada en el entrenamiento de la red generativa.

El resultado de este procedimiento fue un conjunto de máscaras binarias limpias que destacan únicamente las zonas de enfermedad, y una serie de imágenes que contienen sólo los efectos patológicos extraídos del aguacate cabe recalcar que por cada aguacate enfermo se extrajo una máscara como se destaca en la Figura 4 y Figura 5.



**Figura 4.** Resultado del algoritmo OpenCV Scab para detección de Scab: (a) Fruto original; (b) Región segmentada.

Fuente: Los Autores.



**Figura 5.** Resultado del algoritmo OpenCV Scab para detección de Anthracnose: (a) Fruto original; (b) Región segmentada.

Fuente: Los Autores.

Esta perspectiva permitió, además de mantener las propiedades naturales del aguacate saludable, registrar con exactitud las anomalías inherentes a las enfermedades foliares. Posteriormente, estos efectos aislados se usaron como base de datos para entrenar la red generativa, con el objetivo de replicar esos patrones en imágenes sintéticas nuevas.

### 3.2 Generación de Defectos Sintéticos con WGAN

Después de conseguir las máscaras que reflejan únicamente los efectos patológicos de las enfermedades Anthracnose y Scab, se entrenó una red generativa adversaria con el propósito de crear nuevos defectos que replican visualmente las lesiones vistas en los frutos reales. A este fin, se utilizó una Wasserstein GAN con penalización de gradiente (WGAN-GP) porque tiene una habilidad superior para estabilizar el proceso de capacitación y producir imágenes más verosímiles.

#### Aumento de datos previo al entrenamiento

Las imágenes de los defectos extraídos se sometieron a una etapa de aumento de datos (data augmentation) para ampliar la diversidad y solidez del conjunto de entrenamiento previo. Este procedimiento comprendió:

- Rotaciones en 0°, 90°, 180° y 270°.
- Variación aleatoria de brillo.
- Espejado horizontal.

Cada imagen resultante fue redimensionada a 128×128 píxeles y normalizada al rango [-1, 1], lo que resultó en un conjunto suficientemente amplio para entrenar el generador.

#### Arquitectura del modelo

El modelo generativo fue compuesto por:

- Un generador, diseñado con capas Dense, Reshape, y múltiples Conv2DTranspose con activación tanh, el cual recibe vectores de ruido de dimensión 100 como entrada.
- Un crítico (discriminador) que actúa como evaluador del realismo, construido con capas Conv2D, LeakyReLU y Flatten, sin activación final para cumplir con la función de pérdida Wasserstein.

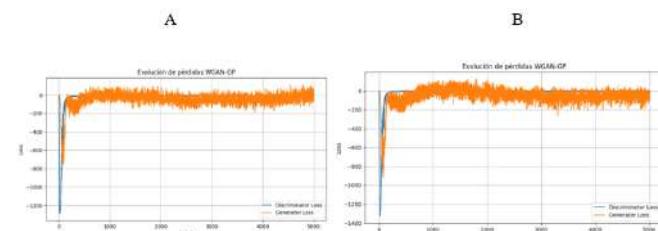
Ambos modelos fueron optimizados con Adam (learning rate 0.0001), configurado con los parámetros recomendados para WGAN-GP ( $\text{beta\_1}=0.5$ ,  $\text{beta\_2}=0.9$ ), y se implementó la penalización de gradiente con  $\lambda = 10$  para reforzar la estabilidad en la retropropagación.

#### Proceso de entrenamiento

El entrenamiento se llevó a cabo durante 5000 épocas, y cada 200 épocas se guardó el modelo con el fin de conservar aquel que ofreciera los mejores resultados, utilizando un tamaño de batch de 32 imágenes. En cada iteración:

- El crítico fue entrenado 5 veces por cada actualización del generador, como sugiere el algoritmo WGAN-GP.
- Se calculó la penalización de gradiente entre muestras reales y generadas, lo que aseguró una convergencia más estable.

A lo largo del entrenamiento se registraron pérdidas en ambos modelos como se muestra en la Figura 6.



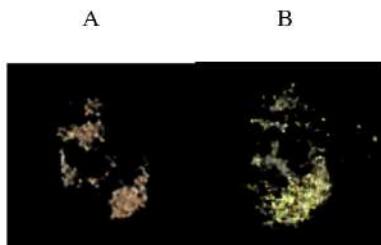
**Figura 6.** Evolución de pérdidas WGAN-GP. (a) Evolución Scab; (b) Evolución Anthracnose.

Fuente: Los Autores.

#### Muestras generadas

Cada cierto número de épocas, se generaron y guardaron imágenes sintéticas de prueba utilizando vectores aleatorios (noise) como entrada al generador entrenado. Estas imágenes fueron visualmente inspeccionadas para verificar la calidad de los defectos generados como se muestra en la Figura 7.





**Figura 7.** Muestras sintéticas de defectos generados por la WGAN-GP. (a) Defecto Scab; (b) Defecto Anthracnose.  
Fuente: Los Autores.

Después de entrenar la WGAN, se produjeron 567 defectos sintéticos nuevos para Scab y 742 para Anthracnose. Se nota en las muestras una coherencia visual con los defectos auténticos obtenidos de las imágenes originales. Esto confirma la habilidad del generador para aprender y reproducir los atributos cromáticos y morfológicos característicos de las enfermedades Scab y Anthracnose.

### 3.3 Síntesis de Imágenes Balanceadas e Inserción de Defectos

Una vez completado el entrenamiento del modelo generativo WGAN-GP, y tras validar la calidad visual de los defectos sintetizados, se procedió a integrarlos dentro de imágenes de aguacates sanos con el fin de construir un conjunto balanceado de clases.

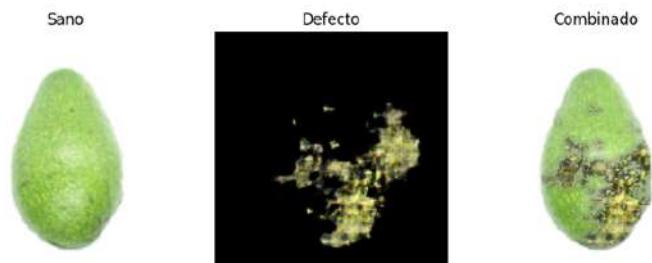
El procedimiento consistió en superponer los efectos patológicos causados en frutas sanas, de modo que se imita la apariencia verdadera de las frutas infectadas por *Scab* y *Anthracnose*. La inserción se llevó a cabo de un modo visualmente realista, teniendo en cuenta la orientación, la proporción y la colorimetría del fruto base. Este método permitió incrementar de manera artificial la clase minoritaria sin duplicar imágenes originales ni modificar la clase *Healthy*.

#### Balanceo del Conjunto de Datos

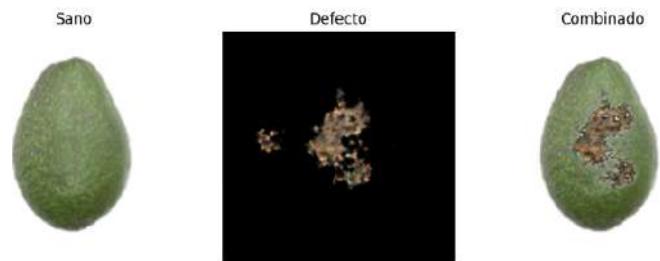
Con el fin de igualar la cantidad de muestras entre clases, se seleccionó una cantidad de 742 imágenes sintéticas para complementar la clase *Anthracnose-E1* y 547 imágenes sintéticas *Scab-E2* respectivamente las cuales se encontraban originalmente subrepresentadas. Las nuevas imágenes fueron distribuidas en sus carpetas correspondientes a su clase *data/Anthracnose-E1* y *data/Scab-E2* para que puedan ser procesadas como parte del flujo estándar de entrenamiento de modelos.

Esta estrategia de inserción permitió transformar un conjunto de

datos originalmente desbalanceado en uno equilibrado en las tres clases, lo cual es esencial para evitar el sesgo de predicción hacia clases mayoritarias durante el entrenamiento del clasificador, en las Figura 8 y Figura 9 se ilustra el proceso de construcción de una muestra sintética:



**Figura 8.** Generación de muestras sintéticas de Anthracnose: fruto original, defecto extraído y resultado combinado.  
Fuente: Los Autores.



**Figura 9.** Generación de muestras sintéticas de Scab: fruto original, defecto extraído y resultado combinado.  
Fuente: Los Autores.

Este procedimiento no solo logró un equilibrio numérico en el conjunto de datos, sino que además amplió la diversidad morfológica de los ejemplos, lo cual le proporcionó al clasificador final una mayor capacidad para generalizar cuando se enfrenta a nuevas muestras.

### 3.4 Comparación de Resultados: Conjunto de Datos Desbalanceados y Balanceados con WGAN

Se entrenaron modelos de clasificación con la misma estructura (InceptionV3 + top model personalizado) en dos contextos diferentes: uno que empleaba el conjunto de datos original desequilibrado y otro que lo hacía con el conjunto equilibrado por medio de la inserción de fallas sintéticas. Esto se hizo para verificar el efecto de balancear los datos conjuntos a través de imágenes sintéticas creadas por la WGAN.

Cabe señalar que los tamaños reportados (3188/795 en el escenario original y 4236/1056 en el balanceado) provienen de aplicar la política de división 80/20 descrita en la sección de Métodos. Esta división fue realizada de manera aleatoria a nivel global, por lo que los conteos específicos por clase pueden variar ligeramente entre las particiones. En la Tabla 3 se resumen los totales de imágenes utilizados en cada escenario, incluyendo el conjunto de prueba independiente de 100 imágenes.

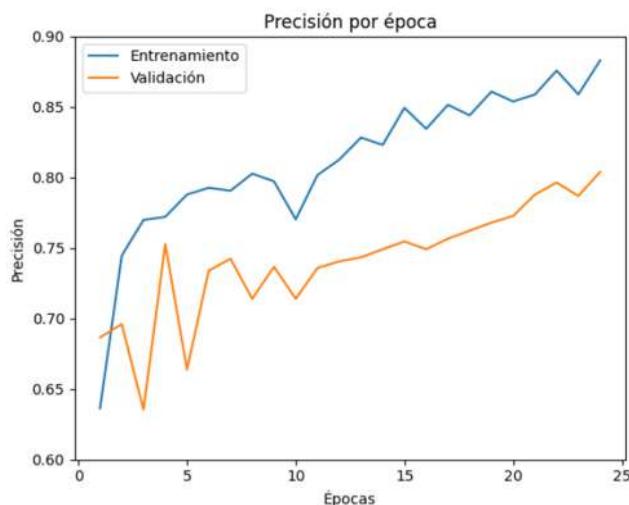
**Tabla 3.** Distribución global de imágenes por partición en los escenarios original y balanceado.

Fuente: Los Autores.

| Escenario  | Entrenamiento | Validación | Total<br>(train+val) | Prueba |
|------------|---------------|------------|----------------------|--------|
| Original   | 3188          | 795        | 3983                 | -      |
| Balanceado | 4236          | 1056       | 5292                 | 100    |

#### Desempeño con Conjunto de Datos Desbalanceado

El entrenamiento con el conjunto original, compuesto por 3188 imágenes de entrenamiento y 795 de validación, mostró un incremento progresivo en la precisión hasta aproximadamente la época 24, momento en el que se alcanzó el mejor valor de validación (83,43%). Apartir de este punto, el modelo se estabilizó y se aplicó EarlyStopping para evitar sobreajuste. La fluctuación en las curvas de pérdida y precisión indica que la distribución desigual de clases influyó en la capacidad de generalización del modelo. Este comportamiento puede observarse en la Figura 10.



**Figura 10.** Desempeño con el Conjunto de Datos Desbalanceado.

Fuente: Los Autores.

#### Desempeño con Conjunto de Datos Balanceado (WGAN)

El modelo que fue entrenado con datos sintéticos compuesto por 4236 imágenes de entrenamiento y 1056 de validación producidos a través de WGAN-GP logró un 97,74 % de precisión

en la validación durante la época 50, en comparación con el 83,43% adquirido en el escenario desequilibrado. Esta mejora de más de 14 puntos porcentuales se debe especialmente a dos factores:

- Disminución de la tendencia hacia la clase mayoritaria: El equilibrio numérico entre las tres clases redujo el sesgo de predicción que se notaba en la situación inicial, lo que permitió al modelo adquirir patrones más representativos de las clases minoritarias.
- Mayor diversidad morfológica: Los fallos producidos por la WGAN-GP aumentaron la gama de formas, texturas y ubicaciones de las lesiones, lo cual mejoró la habilidad del modelo para generalizar en el grupo de validación.

A pesar de que la curva de entrenamiento demostró una baja constante de la pérdida y niveles altos de precisión en validación, el aumento notado debía ser analizado con precaución, pues podría indicar un ajuste a las particularidades propias de los datos sintéticos. Se llevó a cabo un análisis adicional sobre un conjunto de prueba independiente de 100 imágenes que no se habían examinado antes, ni en la validación ni durante el entrenamiento, para abordar esta inquietud.

En este examen, el modelo demostró una notable capacidad de generalización, logrando un macro F1-score de 0.90 y una precisión global del 91%. La matriz de confusión mostró un desempeño ideal en la categoría Healthy (con una precisión y un recall del 100 %), un alto recall en Anthracnose (0,94) y un rendimiento algo inferior en Scab (con 0,92 de precisión y 0,76 de recall), debido a ciertas confusiones con Anthracnose tal como se muestra en la Tabla 4.

**Tabla 4.** Resultados de clasificación en conjunto de prueba independiente (100 imágenes).

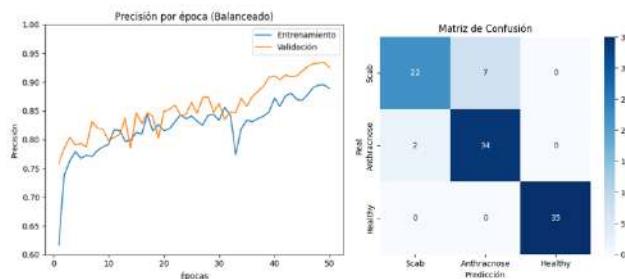
Fuente: Los Autores.

| Clase              | Precisión | Recall | F1-Score | Soporte |
|--------------------|-----------|--------|----------|---------|
| Scab               | 0.92      | 0.76   | 0.83     | 29      |
| Anthracnose        | 0.83      | 0.94   | 0.88     | 36      |
| Healthy            | 1.00      | 1.00   | 1.00     | 35      |
| Promedio Macro     | 0.92      | 0.90   | 0.90     | 100     |
| Promedio Ponderado | 0.91      | 0.91   | 0.91     | 100     |
| Accuracy Global    | -         | -      | 0.91     | 100     |

Estos hallazgos demuestran la efectividad de la estrategia de balanceo por medio de WGAN-GP, como se ilustra en la Figura 11; no obstante, también indican áreas para mejorar, en particular



en lo que respecta a distinguir entre clases con patrones visuales parecidos. Para disminuir la confusión y examinar métodos de clasificación fina (fine-grained classification) que posibiliten diferenciar entre características sutiles de enfermedades que se ven de manera parecida, se aconseja mejorar la producción de defectos en investigaciones futuras.



**Figura 11.** Evaluación del modelo con dataset balanceado: curvas de precisión por época y matriz de confusión final.

Fuente: Los Autores.

### Mejores Épocas de Cada Escenario

**Tabla 5.** Épocas en cada escenario.

Fuente: Los Autores.

| Conjunto de Datos | Mejor Precisión Validación | Época | Pérdida Validación | Tamaño de validación |
|-------------------|----------------------------|-------|--------------------|----------------------|
| Desbalanceado     | 83.43%                     | 24    | 0.4216             | 795                  |
| Balanceado (WGAN) | 97.74%                     | 50    | 0.0704             | 1056                 |

### Validación de la hipótesis mediante bootstrapping

Para evaluar la significancia estadística de la diferencia observada, se aplicó un procedimiento de remuestreo bootstrapping (10 000 iteraciones) sobre las predicciones de validación de ambos escenarios. El análisis arrojó una diferencia media en accuracy de 0,1402 (14,02 puntos porcentuales), con un intervalo de confianza del 95 % de [0,1134, 0,1679] como se muestra en la Tabla 6.

Dado que el límite inferior del intervalo es mayor que cero, se concluye que la mejora es estadísticamente significativa y no atribuible al azar como se observa en la Figura 12.

Si bien este resultado respalda la hipótesis de que el balanceo de clases mediante WGAN-GP mejora el rendimiento del modelo,

**Tabla 6.** Resultados del bootstrapping sobre la exactitud del modelo en el conjunto de prueba.

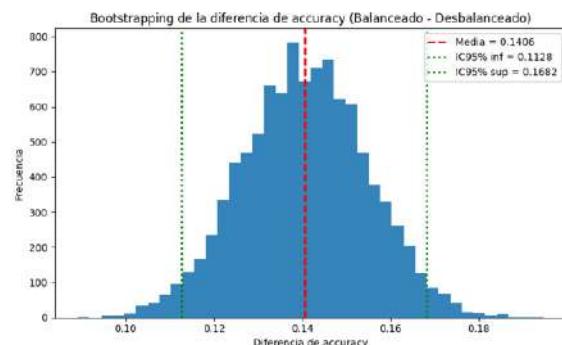
Fuente: Los Autores.

| Parámetro                    | Época        |
|------------------------------|--------------|
| Total de imágenes            | 100          |
| Accuracy observado           | 0.91         |
| Iteraciones (bootstrap)      | 10000        |
| Semilla                      | 2025         |
| IC 95 % (percentil 2.5–97.5) | [0.85, 0.96] |

se reconoce que el bootstrapping no evalúa posibles sesgos derivados de la generación de datos. Por ello, en trabajos futuros se recomienda complementar este análisis con pruebas como McNemar, validación cruzada repetida o evaluación en bases de datos externas, para confirmar la robustez del modelo y descartar efectos de sobreajuste.

### Descripción del procedimiento:

- Se crean vectores binarios que representan aciertos y errores en cada escenario.
- Se realiza un remuestreo aleatorio con reemplazo (bootstrapping) para cada conjunto.
- En cada iteración se calcula la diferencia en accuracy.
- Se construye un intervalo de confianza al 95 % a partir de las diferencias obtenidas.
- Si el límite inferior del intervalo es mayor que cero, la mejora se considera estadísticamente significativa



**Figura 12.** Análisis estadístico bootstrapping: distribución de diferencias de accuracy entre datasets balanceado y desbalanceado.

Fuente: Los Autores.

## Discusión

Los resultados obtenidos respaldan la hipótesis de que el uso de datos sintéticos generados mediante WGAN-GP puede mejorar el rendimiento de modelos de clasificación de imágenes agrícolas. Para evaluar la significancia estadística de la diferencia observada entre el modelo entrenado con el conjunto original y el modelo entrenado con el conjunto balanceado, se aplicó un procedimiento de remuestreo bootstrapping (10 000 iteraciones) implementado en Python. Este análisis arrojó una diferencia media en accuracy de 0,1402 (14,02 puntos porcentuales), con un intervalo de confianza del 95 % de [0,1134, 0,1679], lo que indica que la mejora es estadísticamente significativa y no atribuible al azar.

En el escenario inicial desequilibrado, el modelo llegó a una precisión de validación del 83,43 % en la época 24. Esto ocurrió debido a que se implementó EarlyStopping con una paciencia de 5 épocas, ya que las métricas se estabilizaron. Las fluctuaciones que se aprecian en la curva de validación indican problemas de generalización vinculados con el desbalance. Esto es consistente con lo que Bi et al. (2020) informaron: detectaron aumentos significativos de precisión al emplear WGAN-GP para equilibrar conjuntos de datos agrícolas.

Por otro lado, el modelo que fue entrenado con el conjunto balanceado a través del WGAN-GP logró un 97,74 % de precisión en la validación al llegar a la época número 50, además de una pérdida baja (0,0704). La disminución del sesgo hacia la clase mayoritaria y la inclusión de una diversidad morfológica más amplia en las clases minoritarias pueden ser responsables de este incremento. No obstante, es importante interpretar con precaución las mejoras de esta magnitud, pues podrían ser el resultado de un ajuste a los rasgos particulares de los datos sintéticos. Por esta razón, se aconseja verificar en conjuntos externos e independientes para corroborar la habilidad de generalización.

Estos resultados son coherentes con lo que se ha documentado en la literatura: desde 2017, las GAN han ayudado a incrementar el desempeño de modelos de visión por computadora en labores agrícolas, entre ellas la detección poscosecha, el fenotipado y la clasificación de enfermedades (Carvajal Chávez, 2023; Lu et al., 2022). Investigaciones como las de Bi et al. (2020) han demostrado que la estabilidad y el realismo de las imágenes generadas se incrementan con el uso de variantes como WGAN-GP, que se combinan con técnicas de regularización como cGAN-LSR, logrando así un aumento en la precisión sin menoscabar la robustez.

En resumen, los hallazgos corroboran que WGAN-GP tiene el potencial de ser un instrumento útil para optimizar la actuación de modelos clasificadores en situaciones con clases minoritarias. Sin embargo, para fortalecer esta evidencia, se aconseja que investigaciones posteriores incorporen pruebas de McNemar, validaciones cruzadas que se realicen varias veces y métricas de similitud perceptual (IS, FID). Esto se hace con el objetivo de

evitar sobreajuste y asegurar que el modelo sea generalizable.

## Comparación con métodos clásicos de balanceo

Frente a estrategias tradicionales como el oversampling aleatorio o SMOTE, que duplican o interpolan ejemplos en el espacio de características y pueden inducir sobreajuste o generar instancias poco plausibles a nivel visual, el enfoque propuesto actúa directamente sobre la representación visual del fenómeno patológico. La inserción controlada de defectos segmentados/generados sobre frutos sanos preserva el contexto de iluminación, textura y fondo, a la vez que incrementa la variabilidad morfológica de las lesiones. En contraste, class weights únicamente pondera/ajusta la función de pérdida sin corregir el sesgo en la distribución de muestras. Por su parte, técnicas de mezcla como CutMix, MixUp o Copy-Paste combinan regiones de forma genérica; nuestro pipeline, en cambio, restringe los parches a regiones patológicas plausibles y regula su escala, orientación y colorimetría para mantener coherencia semántica y fotométrica. Como trabajo futuro, una comparación ablation directa contra estas alternativas permitiría cuantificar el beneficio marginal de cada componente del pipeline.

## Calidad y diversidad de los datos sintéticos

Aunque la validación principal se centra en el impacto sobre el clasificador, resulta pertinente auditar los sintéticos con métricas perceptuales como FID (Fréchet Inception Distance) e IS (Inception Score), a nivel global y por clase. El uso de FID/IS ayudaría a detectar colapso de modos, texturas repetitivas o artefactos, y a establecer umbrales mínimos de aceptación antes de incorporar nuevas tandas al entrenamiento.

## Aplicabilidad en condiciones reales de campo

Dado que el conjunto base proviene de un entorno controlado, la transferencia a campo donde cambian iluminación, fondo y occlusiones requiere medidas adicionales: (i) aplicar aumentos fotométricos y geométricos más agresivos durante la síntesis e inserción (variaciones de temperatura de color, sombras, desenfoque, rotaciones y escalas); (ii) introducir domain randomization al posicionar los defectos (rango de tamaños, múltiples lesiones por fruto, ligeras perturbaciones de tono/saturación para compatibilizar con el fondo); y (iii) realizar evaluación externa con imágenes de otras cámaras y ubicaciones, seguida de un ajuste fino del clasificador con un subconjunto anotado en campo. La confusión observada entre *Scab* y *Anthracnose* sugiere explorar enfoques de fine-grained classification y mecanismos de atención local centrados en bordes y texturas de la lesión para mejorar la discriminación en escenarios reales.

## 4. Conclusiones

Esta investigación muestra cómo las redes generativas adversarias, en particular una WGAN que penaliza el gradiente, tienen el potencial de transformarse en un instrumento fundamental para



encarar un desafío común pero complejo: la desproporción entre clases dentro de conjuntos de datos agrícolas. Su puesta en práctica no solamente permitió que la distribución de datos se equilibrara, sino que además los modelos de clasificación basados en aprendizaje profundo mejoraron significativamente su rendimiento. Se mantuvo la coherencia visual global al producir imágenes sintéticas que reflejan con realismo los defectos típicos de enfermedades como el Scab y Anthracnose, lo que resultó en un incremento constante de la estabilidad y precisión del modelo, a su vez un proceso de bootstrapping permitió una firme validación estadística, la cual apoyó este progreso.

La aportación primordial de esta investigación es proporcionar una estrategia que se puede replicar, automatizar y adaptar para enriquecer datos en situaciones donde no hay suficientes muestras disponibles. Este método, a diferencia de las técnicas tradicionales de aumento, no solo homologa la cantidad de ejemplos por clase, sino que también diversifica la representación morfológica, lo cual posibilita que el modelo aprenda patrones más sólidos y transferibles.

Sin embargo, el estudio tiene limitaciones que deben tenerse en cuenta:

- La variabilidad de las condiciones ambientales y de cultivo se ve limitada porque el conjunto de datos proviene de un solo ciclo productivo y una sola localización geográfica, que es México.
- La precisión del clasificador final fue el foco de la evaluación de la WGAN, y no se incluyeron indicadores específicos de similitud perceptual (por ejemplo, IS o FID) que posibiliten medir la diversidad y calidad de las imágenes sintéticas.

Estos descubrimientos son significativos para los desarrolladores e investigadores que se dedican a las imágenes en el ámbito agrícola, particularmente en contextos de agricultura de precisión donde adquirir datos puede ser caro o imposible. Investigaciones futuras podrían ampliar esta metodología a otras enfermedades o cultivos, incluir evaluaciones con conjuntos de datos independientes y emplear métodos avanzados para validar las imágenes producidas. Además, podrían investigar técnicas de regularización que disminuyan el peligro de sobreajuste y arquitecturas generativas más novedosas.

## Agradecimientos

Agradecemos a la ESPAM MFL por su apoyo técnico y académico a través del proyecto institucional Aprendizaje automático y su comportamiento con distintos tipos de conjuntos de datos agropecuarios CUP:91880000.0000.388091 SENPLADES, Ecuador, así como a los responsables del conjunto de datos

utilizados, cuya disponibilidad fue clave para el desarrollo de esta investigación.

## Contribución de los autores

**Luis Jesús Montesdeoca Espinoza:** Conceptualización, Curación de datos, Análisis formal, Metodología, Software, Investigación, Visualización, Redacción – borrador original del artículo, Redacción – revisión y edición del artículo. **Stalin Joel Zambrano Rojas:** Conceptualización, Curación de datos, Análisis formal, Metodología, Software, Investigación, Visualización, Redacción – borrador original del artículo, Redacción – revisión y edición del artículo. **Víctor Joel Pinargote Bravo:** Supervisión, Administración del proyecto, Validación, Redacción – revisión y edición del artículo. **Luis Cristobal Cedeño Valarezo:** Supervisión, Administración del proyecto, Validación, Redacción – revisión y edición del artículo.

## Conflictos de interés

Los autores no tienen conflictos de intereses.

## Referencias bibliográficas

- Bi, W., Wang, X., Xu, W., & Wang, C. (2020). Generation of synthetic training data for deep learning-based plant disease detection system. *Frontiers in Plant Science*, 11, 583438. <https://doi.org/10.3389/fpls.2020.583438>
- Bustamante, M. I., Osorio-Navarro, C., Fernández, Y., Bourret, T. B., Zamorano, A., & Henríquez-Sáez, J. L. (2022). First record of *Colletotrichum anthrisci* causing anthracnose on avocado fruits in Chile. *Pathogens*, 11(10), 1204. <https://doi.org/10.3390/pathogens11101204>
- Carvajal Chávez, C. A. (2023). Uso de técnicas como la regresión y redes neuronales para anticipar el rendimiento del maíz. *Recimundo*, 8(4), 126–135. [https://doi.org/10.26820/recimundo/8.\(4\).diciembre.2024.126-135](https://doi.org/10.26820/recimundo/8.(4).diciembre.2024.126-135)
- Food and Agriculture Organization of the United Nations. (2021). The state of food and agriculture 2021: Making agrifood systems more resilient to shocks and stresses. FAO. <https://doi.org/10.4060/cb4476en>
- Hai, T., Shao, Y., Zhang, X., Yuan, G., Jia, R., Fu, Z., Wu, X., Ge, X., Song, Y., Dong, M., & Yan, S. (2025). An efficient model for leafy vegetable disease detection and segmentation based on few-shot learning framework and prototype attention mechanism. *Plants*, 14(5), 760.

- https://doi.org/10.3390/plants14050760
- Hossen, M. I., Awrangjeb, M., Pan, S., & Al Mamun, A. (2025). Transfer learning in agriculture: A review. *Artificial Intelligence Review*, 58, Article 97. https://doi.org/10.1007/s10462-024-11081-x
- Iparraguirre-Villanueva, O., Guevara-Ponce, V., Paredes, O., Sierra-Liñán, F., Zapata-Paulini, J., & Cabanillas-Carbonell, M. (2022). Convolutional neural networks with transfer learning for pneumonia detection. *International Journal of Advanced Computer Science and Applications*, 13(9), 592–599. https://doi.org/10.14569/IJACSA.2022.0130963
- Lei, L., Yang, Q., Yang, L., Shen, T., Wang, R., & Fu, C. (2024). Deep learning implementation of image segmentation in agricultural applications: A comprehensive review. *Artificial Intelligence Review*, 57, Article 149. https://doi.org/10.1007/s10462-024-10775-6
- Lu, Y., Chen, D., Olaniyi, E., & Huang, Y. (2022). Redes generativas antagónicas (GAN) para el aumento de imágenes en la agricultura: Una revisión sistemática. *Computers and Electronics in Agriculture*, 200, 107208. https://doi.org/10.1016/j.compag.2022.107208
- Paauw, M., Hardeman, G., Taks, N. W., Lambalk, L., Berg, J. A., Pfeilmeier, S., & van den Burg, H. A. (2024). ScAnalyzer: An image processing tool to monitor plant disease symptoms and pathogen spread in *Arabidopsis thaliana* leaves. *Plant Methods*, 20, Article 80. https://doi.org/10.1186/s13007-024-01213-3
- Salem, H. M., Fan, D. P., & Shao, L. (2020). Data augmentation using GAN for improving skin disease classification. *Computer Methods and Programs in Biomedicine*, 198, 105769. https://doi.org/10.1016/j.cmpb.2020.105769
- Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding machine learning: From theory to algorithms. Cambridge University Press.
- Soria Olivas, E., Sánchez-Montaños Isla, A., Gamero Cruz, R., Castillo Caballero, B., & Cano Michalena, P. (2023). Sistemas de aprendizaje automático. RA-MA Editorial.
- Tassi, A., McGough, A. S., & Armitage, D. W. (2022). Applications of generative adversarial networks in agriculture: A review. arXiv. https://doi.org/10.48550/arXiv.2204.04707
- Thayer, A. W., Vargas, A., Castellanos, A. A., Lafon, C. W., McCarl, B. A., Roelke, D. L., Winemiller, K. O., & Lacher, T. E. (2020). Integrating agriculture and ecosystems to find suitable adaptations to climate change. *Climate*, 8(1), 10. https://doi.org/10.3390/cli8010010
- Wang, C., Xia, Y., Xia, L. et al. Dual discriminator GAN-based synthetic crop disease image generation for precise crop disease identification. *Plant Methods* 21, 46 (2025). https://doi.org/10.1186/s13007-025-01361-0
- Wang, S.; Xu, D.; Liang, H.; Bai, Y.; Li, X.; Zhou, J.; Su, C.; Wei, W. Advances in Deep Learning Applications for Plant Disease and Pest Detection: A Review. *Remote Sens.* 2025, 17, 698. https://doi.org/10.3390/rs17040698
- Zhou, Y., Liu, X., & Zhang, J. (2023). A hybrid autoencoder-GAN model for small-scale image synthesis in imbalanced datasets. *Applied Sciences*, 13(2), 905. https://doi.org/10.3390/app13020905





## Monitoreo en tiempo real del ganado: Integración de LoRa, sensores y visualización de datos

## Real-time monitoring of cattle: integration of LoRa, sensors, and data visualization

### Autores

\*Walter Armando Orozco Iguasn<sup>1,2</sup> 

✉ worozco@upse.edu.ec

Juleidy Jailin Tigrero Muñoz<sup>1</sup> 

✉ juleidy.tigrerom@upse.edu.ec

Santiago Andrés Orozco Villarroel<sup>1</sup> 

✉ santiago.orozcovillarroel0058@upse.edu.ec

### Resumen

La producción ganadera en zonas rurales enfrenta importantes desafíos en el monitoreo de la salud y la localización de los animales debido a limitaciones tecnológicas y de conectividad. Para abordar este problema, se desarrolló un sistema de monitoreo en tiempo real basado en comunicación LoRa, diseñado para registrar y visualizar la temperatura corporal y la geolocalización del ganado. La investigación siguió un enfoque cuantitativo, aplicado y de diseño no experimental, validado mediante entrevistas y observaciones con ganaderos de la provincia de Santa Elena, Ecuador, con el fin de asegurar la representatividad del contexto productivo. El sistema implementado integró un Arduino Nano con sensores de temperatura infrarrojos, un módulo GPS y transmisores LoRa, junto con una pasarela basada en ESP8266, una base de datos MySQL y una interfaz gráfica desarrollada en Python. Las pruebas de campo realizadas con cinco reses monitoreadas demostraron una transmisión de datos estable hasta 3 km, confirmando la idoneidad de LoRa para entornos rurales. El análisis estadístico de los datos de temperatura mostró un promedio de 38,4 °C con una variabilidad mínima, mientras que los registros de geolocalización confirmaron una ubicación precisa dentro de los límites establecidos, sin detectar valores atípicos. Además, el mecanismo de alertas mediante la plataforma Blynk permitió la detección oportuna de anomalías de salud y eventos fuera de zona. Los resultados se alinean con las tendencias reportadas por Navia et al. (2024), reforzando la adopción global de sistemas de monitoreo de largo alcance y bajo costo. Sin embargo, a diferencia de enfoques más complejos, este estudio demuestra que una solución simplificada enfocada en la temperatura y la geolocalización puede aportar un valor significativo dentro de las limitaciones tecnológicas y económicas del entorno rural ecuatoriano. Los trabajos futuros se centrarán en la incorporación de parámetros fisiológicos adicionales, fuentes de energía renovables y analítica avanzada para mejorar la escalabilidad y sostenibilidad del sistema.

**Palabras clave:** Internet de las cosas; LoRa; Monitoreo Animal, Ganadería

### Abstract

Livestock production in rural areas faces major challenges in monitoring animal health and location due to technological and connectivity limitations. To address this problem, a real-time monitoring system based on LoRa communication was developed, designed to record and visualize cattle body temperature and geolocation. The research followed a quantitative, applied, and non-experimental design, validated through interviews and observations with farmers in the province of Santa Elena, Ecuador, to ensure representativeness of the production context. The implemented system integrated an Arduino Nano with infrared temperature sensors, a GPS module, and LoRa transmitters, together with an ESP8266-based gateway, a MySQL database, and a graphical Python interface. Field tests with five monitored cattle demonstrated stable data transmission up to 3 km, confirming LoRa's suitability for rural environments. Statistical analysis of temperature data showed an average of 38.4 °C with minimal variability, while geolocation records confirmed accurate positioning within the established boundaries, with no outliers detected. In addition, the alert mechanism using the Blynk platform enabled timely detection of health anomalies and out-of-zone events. The findings align with the trends reported by Navia et al. (2024), reinforcing the global adoption of long-range, low-cost monitoring systems. However, unlike more complex approaches, this study shows that a simplified solution focusing on temperature and geolocation can provide significant value under the technological and economic constraints of rural Ecuador. Future work will focus on incorporating additional physiological parameters, renewable energy sources, and advanced analytics to enhance scalability and sustainability.

**Keywords:** Internet of Things; LoRa; Animal monitoring, Livestock farming

Enviado: 19/10/2025  
Aceptado: 05/11/2025  
Publicado: 07/11/2025



## 1. Introduction

Stockbreeding in Ecuador is one of the most important economic activities in the agricultural sector, as it contributes significantly to the supply of meat, milk, and dairy products, as well as generating employment in rural areas and boosting the local economy. This activity is carried out in various regions of the country, taking advantage of the climatic and geographical diversity that allows for the breeding of different breeds and production methods (Jaramillo y Ríos, 2025). Like any activity, it faces challenges related to productivity, health management, and environmental sustainability. In this context, the use of modern technologies, such as monitoring systems, sensors, artificial intelligence, and data management platforms, is a key tool for improving efficiency, ensuring animal welfare, and boosting the competitiveness of Ecuadorian livestock farming (Tzanidakis et al., 2023).

Telecommunications systems play an important role in technological development across all sectors. Livestock farming, as a driver of rural development, faces significant challenges in incorporating tools that support aspects such as monitoring the health and location of livestock, among other issues (Ahmed y Gallardo, 2022); due to a lack of technological infrastructure and limited connectivity in remote areas. In this context, low-power wide-area network (LPWAN) technologies have emerged as promising solutions for smart agriculture, offering long-distance communication with low energy consumption, ideal for areas with poor cellular coverage (Jaramillo y Ríos, 2025).

The incorporation of technology in the livestock sector is developing greatly, influenced not only by the availability of equipment but also by the medium- and long-term benefits it brings to livestock production rates. There are many areas and processes in which technology is and would be a differentiating factor for continuous improvement (Xiao et al., 2022). Tasks such as environmental monitoring, real-time tracking, detection of atypical patterns, welfare, safety, and operational efficiency are some examples that require attention (Schulthess et al., 2024).

As part of the modernization of the livestock sector, monitoring critical variables such as the geographical location and body temperature of livestock is essential to ensure productivity and animal welfare (Montalván et al., 2024). The ability to track cattle movements in real time allows for optimized pasture management, loss prevention, and improved herd safety, while continuous monitoring of body temperature provides an early indicator of disease, heat stress, or physiological changes that directly affect health and productive performance (Ojo et al., 2022).

(Navia et al., 2024) conducted a systematic literature review on sensor-based systems and applications for livestock vital signs monitoring, applying the PRISMA methodology to identify trends and research gaps. Their study analyzed 21 scientific papers published between 2017 and 2023, as well as several commercial products. The results highlighted that the most widely used communication technologies are long-range ones, especially LoRaWAN, followed by WiFi and GSM-based solutions. Most of the reviewed proposals targeted cattle, employing fully mobile devices such as collars or ear tags, which enable monitoring in grazing and large-scale environments. In terms of monitored parameters, the predominant variables were geolocation, movement, and body temperature, while other vital signs such as heart rate, respiration, blood pressure, or oxygen saturation received less attention.

This background study supports the relevance of the present research, as it aligns with the trend of using LoRa-based communication for long-range monitoring, while addressing specific challenges of cattle management in rural Ecuadorian contexts. From this perspective, this study proposes a monitoring system that integrates both variables using LoRa communication, facilitating timely decision-making and promoting more efficient, sustainable, and innovative livestock management. This system seeks to generate alerts in the event of potential health risks or losses, facilitating timely intervention by the personnel in charge. The choice of LoRa technology is due to its ability to transmit data over long distances with low energy consumption, ideal for this type of environment (Rivera Guzmán et al., 2022).

The system design includes a modular structure consisting of collars equipped with sensors, a database for storing information, and a graphical interface that allows users to view readings for each animal. The solution represents a significant step forward in the modernization of agricultural practices, aiming to reduce the margin for human error and optimize response times to incidents, contributing to the operational efficiency of farms and the sustainability of local livestock production. The project is positioned as a viable, scalable technological solution aligned with the needs of the Ecuadorian countryside today.

## 2. Materials and Methods

The methodology of this research was designed to guarantee the validation of the proposed cattle monitoring system under real production conditions. A quantitative, applied, and non-experimental design was followed, structured in four phases:

Problem analysis: To identify the main challenges in livestock management, semi-structured interviews were conducted

with local farmers and technicians, complemented by direct observation on several farms in the province of Santa Elena, Ecuador. This triangulation of techniques allowed the identification of technological limitations, environmental conditions, and infrastructure constraints. Based on this analysis, a representative case study farm was selected, with a total population of 70 cows, of which 5 were randomly chosen for daily monitoring during the experimental period.

**Prototype development:** Prototype monitoring was developed, consisting of collars equipped with infrared temperature sensors (MLX90614), a GPS module (Ublox Neo), and an Arduino Nano microcontroller connected to a LoRa SX1278 module for long-range wireless communication. A NodeMCU ESP8266 module was used as the receiver to store data on a local server. The design schematics were created in Proteus to validate the connections and ensure system stability.

**Software design:** The waterfall model was adopted for the development of software components. The system architecture comprised three layers: the graphical user interface, developed in Python with CustomTkinter; the backend processing module, also implemented in Python, responsible for receiving and processing sensor data via TCP/IP; and a MySQL database for storage and historical queries.

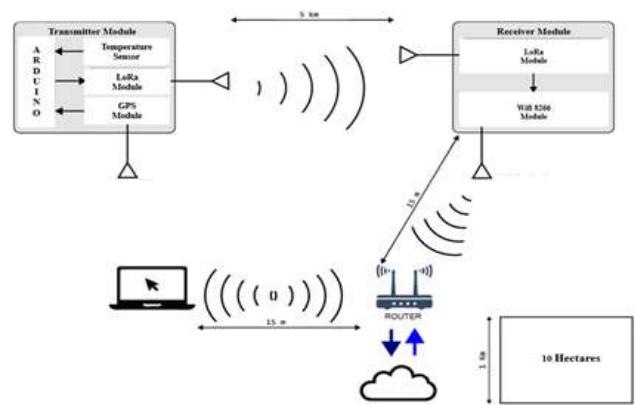
**Field testing:** Field tests were carried out on the selected farm over a period of four weeks. The monitoring collars transmitted temperature and geolocation data in real time to the LoRa receiver, which sent the information to the database. Data were analyzed to evaluate system performance in terms of communication stability, operational coverage, accuracy of location, and reliability of temperature readings. Additionally, an alert system was implemented using the Blynk platform to send automatic notifications by email when anomalies (thermal variations or out-of-zone movements) were detected.

### 3. Results and Discussion

The research includes architecture that uses innovative technologies capable of intelligently processing the measured variables. The results achieved are the Monitoring Equipment and Control Software.

#### 3.1. Monitoring Equipment

The equipment consists of several components; Figure 1 shows the block diagram of the equipment; the transmitting device located on the cattle will be responsible for sending temperature and geolocation data to the receiver, while the receiver stores the data on a data server.



**Figure 1.** Block diagram: Transmitter and Receiver

Source: The authors

The materials and equipment used were evaluated and selected based on the environment where monitoring takes place. Two different microcontrollers are used, one for the transmitter and one for the receiver. The main components of the equipment are microcontrollers, GPS module, temperature sensor, LoRa module, and a voltage regulator.

**Microcontroller:** Arduino Nano ATmega328, used for reading the sensors, small size, low power consumption, and ease of programming. It is ideal for portable applications, such as cattle collars, as it can be installed in small spaces without compromising functionality (Taryana, 2021).

**GPS, Ublox Neo,** a module chosen for its cost-effectiveness, ease of use, and location accuracy, which allows real-time tracking of livestock, essential for identifying unusual movements. It also has an external antenna that improves reception in rural areas.

**LORA module:** SX1278, communication technology characterized by its wide range, which can reach up to 5 km in open field, and low energy consumption. The SX1278 module allows portable devices, such as collars, to send data to the central node without data loss.

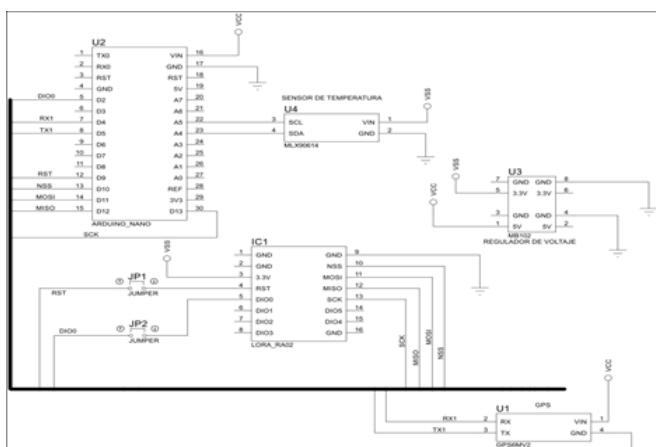
**Gy-906 MLX90614 Infrared Temperature Sensor:** a sensor that allows the body temperature of animals to be measured without direct contact, helping to prevent discomfort or infection in livestock. Among its features is that it uses the I2C protocol to establish communication with the microcontroller.

**Power supply module mb102:** This module was chosen because the microcontrollers and components have different voltage outputs of 5V and 3.3V; therefore, a voltage regulator was selected to ensure that the outputs are compatible with all the sensors and microcontrollers in the system.

**Receiver microcontroller:** A different microcontroller is used for the receiver device than for the transmitter, in this case the NODE MCU ESP8266. Among the other components used for data reception are the same LoRa module and voltage power supply module that were used in the transmitter device.



The diagram of the transmitter solution is shown in Figure 2 and that of the receiver in Figure 3. The diagrams show the connections of all the system components: microcontroller, GPS module, temperature sensor, LoRa module, and power supply. The designs were created in Proteus, which allows us to validate and ensure that each connection is correct, orderly, and error-free.

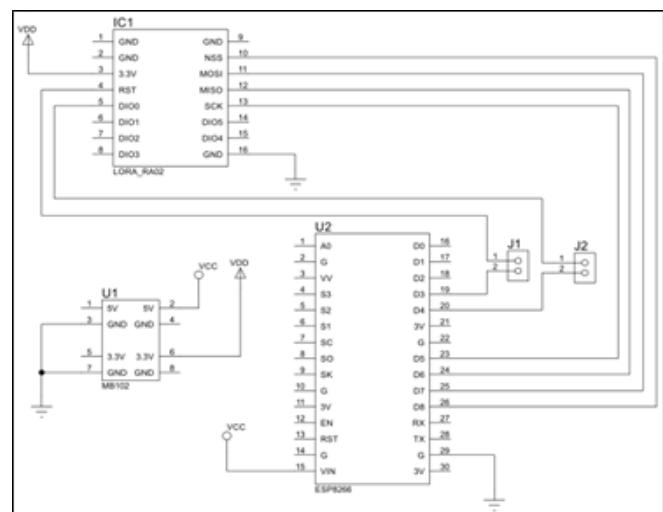


**Figure 2.** Transmitter Schematic Design  
Source: The authors

In order for the sensors and the microcontroller (Arduino Nano) to communicate with each other, several protocols were implemented, each with different characteristics. The MLX90614 infrared temperature sensor uses the I2C protocol, which makes it easy to connect several devices with only two data cables (SCL and SDA). In contrast, the NEO-6M GPS module communicates via the UART protocol, which is asynchronous and ideal for sending location data. Finally, the LoRa SX1278 module uses the SPI protocol, which allows data to be sent faster and is very useful for the microcontroller to communicate with the radio frequency modules. The TCP/IP protocol is also used to send information from the receiving device to the server. This protocol ensures secure communication, making it possible for the data collected by the collars, both position and temperature, to reach the control system without loss of information (Harini Kolamunna, 2016).

### 3.2. Control Software

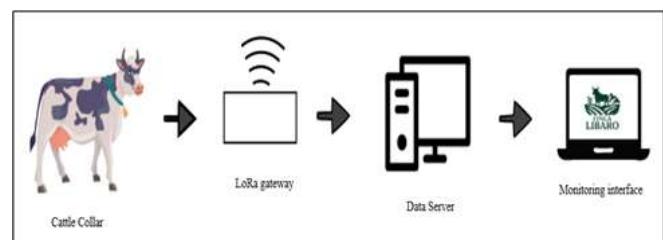
The waterfall model was used as a reference for the development of the system. This model rigorously orders the stages of the software development process, such that the start of each stage must wait for the completion of the previous stage (Rivera Guzmán et al., 2022). It includes phases such as requirements



**Figure 3.** Receiver Schematic Design  
Source: The authors

analysis, system design, implementation, testing, deployment, and maintenance. In each of the phases, the functional requirements that enable the project's objective to be met have been identified (Maida y Pacienzia, 2015).

The software architecture is based on three elements: the user interface (frontend), the processing logic (backend), and the database. Python with the CustomTkinter library was used for the frontend, which allowed us to create a graphical desktop interface that is intuitive for the end user. The backend, also developed in Python, acts as a bridge between the interface and the database. This module contains the business logic that processes data from IoT devices, which arrive via HTTP requests using the TCP/IP protocol. All the information collected is then stored in a MySQL database. This database organizes records on cattle and also facilitates queries and historical analysis.

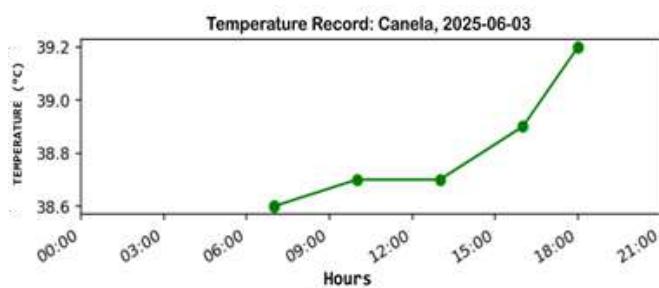


**Figure 4.** Communication scheme  
Source: The authors

Figure 4 shows the integrated solution, composed of: i) Cattle Collar, with Arduino, temperature sensor, GPS, and LoRa for long-range transmission; ii) LoRa Gateway (receiver), with LoRa and ESP8266 modules that relay data to a local server; iii) Data Server, which stores and processes the information in a MySQL database; and iv) Monitoring Interface (Python), a graphical system that displays livestock location, temperature, and alerts in real time or historically. The tests carried out have allowed us to obtain data associated with the variables: temperature, location and Out-of-zone alerts.

Temperature: Figure 5 shows the body temperatures recorded for the cow "Canela" on June 7, 2025. There is an upward trend in body temperature throughout the day, starting with values of 38.2°C at 7:00 a.m. and 9:00 a.m. and reaching a maximum of 38.9°C at 5:00 p.m. This progressive increase could be due to environmental factors such as midday heat or the animal's physical activity. It is important to note an anomaly at 1:00 p.m., where the temperature drops to 37.7°C, which may be due to an atypical reading or a brief pause in the animal's activity.

On the other hand, the most relevant data in the table in Figure 5 is that at 7:00 a.m., the cattle were recorded as being outside the monitoring area. This event could imply a possible risk of escape. However, for the rest of the day, Canela remained within the defined area. This type of information is essential for cattle management, as continuous monitoring allows for the anticipation of possible health problems and timely decisions on the farm.



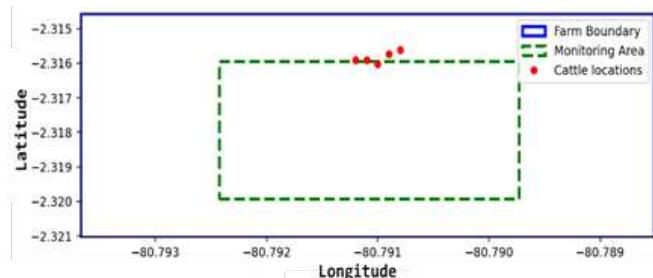
**Figure 5.** Temperature recorded by hour of the day  
Source: The authors

Location: Figure 6 shows the route taken by the cow "Estrella" on June 8, 2025. Three elements stand out in the graph: the total boundary of the farm in blue line, the monitoring area established in green dotted line, and the recorded location points of the cow in red dots. All location records are within the monitoring area, indicating that the monitoring system is adequately fulfilling its tracking function within the defined operating limits. This ensures the control and safety of the animal without the need for manual location methods.

The location records were analyzed to determine spatial distribution within the monitoring area. All points fell within the defined boundary, confirming the accuracy of the GPS module. The concentration of positions in a reduced zone suggests limited

mobility during the observation period, a pattern consistent with grazing behavior. From a statistical perspective, no outliers were detected in the dataset, which validates the stability of the geolocation system

In addition, the concentration of points in a specific area suggests that the cattle have remained within a limited range of movement during the observation period. This may be due to factors such as the availability of food or shade in that area. This information is crucial for farm managers, as it allows them to evaluate behavior patterns, make decisions about grazing rotation, or detect possible anomalies in cattle activity. Overall, route analysis supports the efficiency of the system and contributes to improving the productive and sanitary management of the cattle.



**Figure 6.** Location of cattle  
Source: The authors

Out-of-zone alerts: The alert system uses the Blynk platform, together with an ESP8266 board, to send automatic email notifications. Sensors provide real-time data on each animal's body temperature and GPS location. If the temperature exceeds a threshold, the system generates an immediate alert that is sent by email, including the time, latitude, longitude, and exact temperature of the cattle. A similar process occurs when the cattle are outside the established limits, triggering an "Out of zone" alert.

The results demonstrate the effectiveness of the proposed system by jointly integrating body temperature and geolocation monitoring of cattle. The recorded thermal variations made it possible to identify patterns associated with environmental factors as well as potential anomalies in measurements, while location tracking confirmed the system's ability to keep animals controlled within defined boundaries. Furthermore, the automatic alert mechanism enhances responsiveness to health or escape risks, reducing reliance on manual supervision. Overall, these findings indicate that continuous monitoring through low-cost technologies such as LoRa and ESP8266 constitutes an effective tool to improve herd management and ensure animal welfare.

The results of this study align with the systematic review by (Navia et al., 2024) which highlights body temperature and geolocation as key parameters for livestock monitoring and confirms the predominance of long-range technologies such as LoRaWAN. The stable communication range of up to 3 km and the use of mobile collars in this research reinforce these global



trends, demonstrating the feasibility of portable, animal-worn solutions in open grazing conditions. However, unlike many systems reviewed by Navia et al. that integrate multiple vital and behavioral signs or propose renewable energy options, the present prototype focused solely on temperature and geolocation, prioritizing simplicity, low cost, and adaptability to rural Ecuadorian contexts, while leaving future opportunities for expansion and energy autonomy.

Overall, the findings not only validate the global trend toward LoRa-based livestock monitoring but also contribute new evidence on its practical viability in Ecuadorian production systems, where cost, simplicity, and adaptability to connectivity limitations are decisive for adoption.

#### 4. Conclusions

The development of the monitoring system enabled real-time management of cattle positioning and temperature. LoRa technology proved effective in maintaining stable communication between the transmitting nodes and the central server, even in areas with limited coverage, which supports its suitability for rural environments. The transmitting and receiving devices designed fulfilled their purpose, correctly integrating GPS and temperature sensors with Arduino Nano and ESP8266 microcontrollers. These devices were able to capture, process, and send the necessary information to the server, which allowed for the establishment of a functional network connecting the cattle to a monitoring center. The LoRa communication network was successfully established, achieving effective operational coverage of up to 3 kilometers during testing. This allowed for a continuous flow of information without data loss. Implementing a solution of this type allows for immediate decisions to be made in the face of potential health risks, unusual movements, and even unsafe situations, demonstrating that technology can be adapted to the needs of the livestock environment, promoting its sustainable development.

This study confirms the relevance of monitoring body temperature and geolocation through LoRa-based communication, validating global findings reported by (Navia et al., 2024). Unlike more complex systems, the proposed simplified solution demonstrates practical value and feasibility under the technological and economic constraints of rural Ecuador. Future developments of this research may include the integration of additional physiological parameters, such as heart rate or respiration, to provide a more comprehensive view of cattle health. Likewise, incorporating renewable energy sources, such as solar panels, could improve device autonomy and reduce maintenance

needs. Finally, the integration of advanced data analytics or machine learning techniques could enhance anomaly detection and decision-making, further strengthening the role of IoT in sustainable livestock management

#### Contributions of the Authors

**Walter Armando Orozco Iguasnia:** Conceptualization, Methodology, Writing-revision and editing paper. **Juleidy Jailin Tigrero Muñoz:** Research, Data curation, Formal analysis and Software. **Santiago Andrés Orozco Villarroel:** Original draft paper.

#### Conflict of interest

The authors declare that they have no conflict of interest.

#### References

- Ahmed, M., & Gallardo, L. (2022). LoRa Based IoT Platform for Remote Monitoring of Large-Scale Agriculture Farms in Chile. *Sensor*, 22(8), 2824. <https://doi.org/10.3390/s22082824>
- Harini Kolamunna, J. C. (2016). Are wearable devices ready for HTTPS? Measuring the cost of secure communication protocols on wearable devices. *arXiv*, 2, 9. <https://doi.org/10.48550/arXiv.1608.04180>
- Jaramillo, Y., & Ríos, P. (2025). Tendencias Tecnológicas con IoT en la Ganadería 4.0 Aplicables en Ecuador. *Ciencia Latina Revista Científica Multidisciplinar*, 8(6), 5947-5974. [https://doi.org/10.37811/cl\\_rcm.v8i6.15294](https://doi.org/10.37811/cl_rcm.v8i6.15294)
- Maida, E. G., & Pacienzia, J. (2015). Metodologías de desarrollo de software.
- Montalván, S., Arcos, P., & Sarzosa, P. (2024). Technologies and Solutions for Cattle Tracking: A Review of the State of the Art. *Sensors*, 24(19).
- Navia, M., Chancay-García, L., Cedeño-Sarmiento, C., & Mendoza-Briones, D. (2024). A Systematic Literature Review about Sensors Systems and Applications for Livestock Vital Signs Monitoring. *Revista Politécnica*, 69-78.
- Ojo, M. O., Viola, I., Baratta, M., & Giordano, S. (2022). Practical Experiences of a Smart Livestock Location Monitoring System Leveraging GNSS, LoRaWAN and

- Cloud Services. *Sensor*, 22(1). <https://doi.org/https://doi.org/10.3390/s22010273>
- Rivera Guzmán, E., Mañay Chochos, E., & Chiliquinga Malliquinga, M. (2022). LoRa Network-Based System for Monitoring the Agricultural Sector in Andean Areas: Case Study Ecuador. *Sensor*, 22(18), 6743. <https://doi.org/https://doi.org/10.3390/s22186743>
- Schulthess, L., Longchamp, F., Vogt, C., & Magno, M. (2024). A Lora-Based and Maintenance-Free Cattle Monitoring System for Alpine Pastures and Remote Locations. *Electrical Engineering and Systems Science*. <https://doi.org/https://doi.org/10.48550/arXiv.2406.06245>
- Taryana, S. (2021). Antarmuka ublox NEO-6M GPS Module dengan NodeMCU ESP8266.
- Tzanidakis, C., Tzamaloukas, O., Simitzis, P., & Panagakis, P. (2023). Precision Livestock Farming Applications (PLF) for Grazing Animals. *Agriculture*, 13(2), 288. <https://doi.org/https://doi.org/10.3390/agriculture13020288>
- Xiao, F., Weizheng, S., & Yanling, Y. (2022). Remote monitoring system for livestock environmental information based on LoRa wireless ad hoc network technology. *International Journal of Agricultural & Biological Engineering*, 15(4). <https://doi.org/10.25165/j.ijabe.20221504.6708>





## Desarrollo de un Chatbot con Inteligencia Artificial para orientación vocacional según el perfil competencial de estudiantes de tercero bachillerato

### Development of a Chatbot with Artificial Intelligence for Career Guidance Based on the Competency Profile of Third-Year High School Students

#### Autores

\*Fernando Ericsson Montoya Naguas 

 fmontoya1@utmachala.edu.ec

Carlos Elian Lalangui Flores 

 clalangui1@utmachala.edu.ec

Fausto Redrován Castillo 

 fredrovan@utmachala.edu.ec

Freddy Aníbal Jumbo Castillo 

 ffjumbo@utmachala.edu.ec

Universidad Técnica de Machala,  
Facultad de Ingeniería Civil, Carrera de  
Tecnologías de la Información, Machala,  
El Oro, Ecuador.

\*Autor para correspondencia

#### Comó citar el artículo:

Montoya Naguas, F.E., Lalangui Flores, C.E., Redrován Castillo, F.F. & Jumbo Castillo, F.A. (2025). Desarrollo de un Chatbot con Inteligencia Artificial para orientación vocacional según el perfil competencial de estudiantes de tercero bachillerato, *Informática y Sistemas* 9(2), pp. 184-196. <https://doi.org/10.33936/isrtic.v9i2.7907>

Enviado: 27/09/2025

Aceptado: 11/11/2025

Publicado: 18/11/2025

#### Resumen

Este artículo presenta el diseño e implementación de un chatbot con inteligencia artificial creado para apoyar la orientación vocacional de estudiantes de tercer año de bachillerato en Ecuador, la iniciativa responde ante la escasez de especialistas en orientación vocacional y la escasa disponibilidad de plataformas tecnológicas que faciliten este proceso en los colegios. El sistema ofrece sugerencias ajustadas al perfil de cada estudiante y a la oferta académica de la Universidad Técnica de Machala (UTMACH), empleando el modelo de lenguaje Deepseek y técnicas de ingeniería en prompts para interpretar respuestas abiertas y valorar sus competencias de acuerdo con el currículo priorizado del Ministerio de Educación. El desarrollo del chatbot se realizó siguiendo la metodología ágil Scrum, para evaluar su efectividad, se aplicó un estudio cuasiexperimental con 351 estudiantes, cuyos resultados fueron contrastados con el test vocacional de la SENESCYT. La comparación evidenció una coincidencia del 80.1% entre ambos instrumentos y una clara preferencia de los estudiantes por el chatbot: el 76% valoró positivamente la interacción conversacional, la claridad de las preguntas y la pertinencia de las recomendaciones, destacando además la accesibilidad de la herramienta. De acuerdo con la hipótesis planteada, los hallazgos confirman que el Chatbot puede emplearse como alternativa válida al test SENESCYT, complementando los procesos tradicionales de orientación vocacional. Aunque se reconocen limitaciones como el muestreo no probabilístico y la ausencia de aprendizaje continuo en el modelo de IA, la herramienta demuestra un potencial significativo para democratizar el acceso a recomendaciones vocacionales personalizadas en el contexto educativo ecuatoriano.

**Palabras clave:** Orientación vocacional; Inteligencia artificial; Chatbot; Educación; Competencias.

#### Abstract

This article presents the design and implementation of an artificial intelligence-based chatbot developed to provide personalized career guidance to senior high school students in Ecuador. The tool addresses the limited availability of vocational guidance professionals and the lack of interactive technological platforms in educational institutions, offering recommendations aligned with students' individual competencies and the academic programs of the Technical University of Machala (UTMACH). The system leverages the Deepseek language model and prompt-engineering strategies to process open-ended responses, assessing mathematical, digital, communication, and socio-emotional skills in accordance with the prioritized curriculum established by the Ministry of Education. The chatbot was developed following the agile Scrum methodology, and its effectiveness was evaluated through a quasi-experimental study with 351 participants. The recommendations generated by the chatbot were compared with those of the official SENESCYT vocational test. Results revealed an 80.1% agreement between the two methods, highlighting the chatbot's clarity, empathy, and perceived usefulness. Furthermore, 76% of the students expressed a preference for the AI-driven system, valuing in particular the conversational interaction, accessibility, and relevance of the recommendations. In line with the stated hypothesis, the findings confirm that the chatbot can serve as a valid alternative to the SENESCYT test, complementing traditional vocational guidance practices. Although limitations such as non-probability sampling and the absence of continuous learning in the AI model are acknowledged, the system demonstrates significant potential to democratize access to personalized career recommendations within the Ecuadorian educational context.

**Keywords:** Vocational Guidance; Artificial Intelligence; Chatbot; Education; Competencies.



## 1. Introduction

Durante los últimos años, la orientación vocacional ha emergido como una dificultad para los sistemas educativos, especialmente en países como Ecuador, donde muchos estudiantes de bachillerato afrontan decisiones sobre su desarrollo académico sin contar con el acompañamiento necesario. La limitada disponibilidad de profesionales con formación suficiente en orientación vocacional y la falta de herramientas tecnológicas efectivas en los centros educativos, como software educativo, plataformas interactivas y sistemas de evaluación de intereses y habilidades, restringen el acceso a un asesoramiento adecuado, lo que puede derivar en decisiones desinformadas, abandono de carreras y baja satisfacción profesional (Erazo & Rosero, 2021; Torres-Sevillano et al., 2025).

Actualmente se conoce varios avances sobre la inteligencia artificial (IA), esto ha permitido abrir nuevas posibilidades en el ámbito educativo, debido a que existen herramientas como el procesamiento de lenguaje natural (PLN) y el aprendizaje automático, gracias a esto se ha facilitado el desarrollo de chatbots capaces de simular conversaciones humanas y recopilar información en tiempo real para ofrecer respuestas personalizadas a cada usuario (Caldarini et al., 2022; Debets et al., 2025; Phuttwamong & Chatwattana, 2025). Además, estas herramientas representan un recurso valioso para ampliar el acceso a la orientación vocacional debido a que permite acompañar a un gran número de estudiantes sin depender de la presencia constante de un orientador (Davar et al., 2025). De igual manera, investigaciones recientes señalan que los chatbots educativos no solo cumplen con la función del apoyo conversacional, sino que también favorecen la autorregulación y el aprendizaje autónomo de los estudiantes (Haristiani et al., 2022; Labadze et al., 2023; Lee et al., 2025; Lucana & Roldán, 2023; Okonkwo & Ade-Ibijola, 2021).

Actualmente, el sistema educativo ecuatoriano utiliza como herramienta principal el test vocacional de la SENESCYT, una prueba estática con afirmaciones cerradas que requiere interpretación profesional para emitir diagnóstico (Ministerio de Educación del Ecuador, 2025). Si bien ha sido útil, su aplicación masiva sin retroalimentación inmediata limita la personalización del proceso. Como alternativa, se realizó una investigación en la Universidad Técnica de Machala (UTMACH) para el desarrollo de un test vocacional basado en aptitudes con elementos de IA, pero carece de interacción dinámica (Mayon & Ruiz, 2024), y hasta la actualidad no se ha aplicado.

En este contexto, el presente estudio introduce un enfoque innovador al proponer un chatbot vocacional inteligente que no solo reproduce procesos de recomendación automatizados, sino

que incorpora una capa de análisis contextual y conversacional, a diferencia de los sistemas previos, que se limitan a la clasificación de respuestas cerradas o análisis de patrones, el chatbot desarrollado combina las competencias del Currículo Nacional Priorizado con los valores institucionales JIS (Justicia, Innovación y Solidaridad), generando recomendaciones más interpretativas y personalizadas, dado que este enfoque busca demostrar que la orientación vocacional puede ser asistida por IA sin sustituir la reflexión humana, sino potenciándola mediante diálogo y análisis.

Hasta la fecha, no se han documentado estudios en Ecuador que comparan directamente la efectividad de un chatbot vocacional con el test tradicional de la SENESCYT, lo que resalta la oportunidad de evaluar la pertinencia, coherencia y aceptación de las recomendaciones generadas por sistemas basados en inteligencia artificial en programas de orientación vocacional (Jaramillo, 2024). Estudios internacionales ya han explorado esta línea, comparando la efectividad de chatbots frente a evaluaciones tradicionales en contextos educativos (Toscu, 2024).

Frente a estas limitaciones, se propone como alternativa el desarrollo de un chatbot vocacional inteligente capaz de interactuar de manera directa con los estudiantes mediante lenguaje natural. A diferencia de las pruebas tradicionales, esta herramienta recopila información abierta relacionada con intereses, pasatiempos, aspiraciones y habilidades, lo que le permite valorar competencias comunicativas, matemáticas, digitales y socioemocionales de acuerdo con el currículo priorizado del Bachillerato Ecuatoriano. Asimismo, integra principios vinculados al perfil de egreso estudiantil como la justicia, la innovación y la solidaridad (Ministerio de Educación del Ecuador, 2025), ofreciendo de esta forma recomendaciones más integrales y contextualizadas.

En términos teóricos, esta investigación refuerza la relación entre las competencias evaluadas por el chatbot y los estándares del currículo priorizado, explicando cómo dichas competencias se operacionalizan en las preguntas generadas por IA, por ejemplo, los ítems orientados al razonamiento numérico permiten observar el dominio de la competencia matemática, mientras que las respuestas reflexivas o empáticas revelan dimensiones socioemocionales de esta manera la estructura de análisis posibilita una evaluación más flexible y contextual que los métodos tradicionales de opción múltiple. No obstante, es necesario reconocer las posibles limitaciones éticas y pedagógicas del uso de IA en la orientación vocacional, entre ellas la dependencia tecnológica, la interpretación algorítmica de respuestas humanas y el riesgo de sesgo en las recomendaciones, por tal motivo, el presente trabajo asume un enfoque responsable,

considerando al chatbot como un apoyo complementario al juicio humano, no como un sustituto del orientador educativo.

### 1.1. Hipótesis Central

La implementación de un chatbot basado en inteligencia artificial incide en la precisión y personalización de las recomendaciones vocacionales para la selección de una carrera universitaria en estudiantes de tercero de bachillerato, igual o mejor de cómo lo hace un test estandarizado como sería el de la SENESCYT.

### 1.2. Objetivo General

Diseñar e implementar un chatbot vocacional inteligente alineado con el currículo priorizado con énfasis en competencias del Bachillerato Ecuatoriano y la oferta académica de la Universidad Técnica de Machala (UTMACH).

### 1.3. Objetivos Específicos

**1. Recolectar información clave** sobre la orientación vocacional en los colegios del país y la formación escolar basada en competencias, con el fin de establecer una base teórica y contextual para el desarrollo del chatbot.

**2. Identificar los requerimientos funcionales y técnicos** necesarios para la creación del chatbot, tomando como referencia el currículo nacional priorizado con énfasis en competencias.

**3. Codificar y desarrollar el chatbot vocacional** conforme a los requerimientos establecidos, garantizando su correcta integración con los objetivos del estudio.

**4. Evaluar la pertinencia y coherencia** de las recomendaciones emitidas por el chatbot en comparación con el test vocacional tradicional aplicado por la SENESCYT.

**5. Analizar la percepción de los estudiantes** sobre la utilidad, claridad y aceptación del sistema como herramienta de apoyo en la toma de decisiones académicas.

### 1.4. Descripción del Estudio

El estudio se orienta al diseño, desarrollo e implementación de este chatbot, alineado con la oferta académica de la Universidad Técnica de Machala (UTMACH), y a la evaluación de su efectividad frente al test vocacional de la SENESCYT. Para ello, se recurrió a un diseño cuasiexperimental con la participación de 351 estudiantes de tercer año de bachillerato, analizando tanto la coherencia de las recomendaciones emitidas como la percepción de los usuarios respecto a la utilidad y funcionalidad del sistema. Esta metodología permitió valorar la pertinencia de la herramienta en comparación con el test tradicional, considerando aspectos de validez y aceptación estudiantil.

A partir de ello, surgió la siguiente pregunta de investigación:

¿Puede un sistema conversacional basado en inteligencia artificial generar recomendaciones vocacionales pertinentes y aceptadas por los estudiantes, comparables con lo ofrecido por un test tradicional?

## 2. Materiales y Métodos

La investigación se diseñó bajo un enfoque mixto que combinó técnicas cualitativas y cuantitativas, orientadas a examinar la funcionalidad técnica del chatbot vocacional y, al mismo tiempo, la aceptación que generó entre los estudiantes. El estudio es de tipo correlacional descriptivo, al centrarse en establecer el grado de coincidencia entre las recomendaciones generadas por el sistema conversacional y las del test tradicional, así como describir la percepción del usuario.

La investigación se aplicó en tres instituciones educativas: Unidad Educativa Particular “Hermano Miguel” (Machala, El Oro), Unidad Educativa Particular “Divino Niño” y la Unidad Educativa “Ponce Enríquez”, ambas ubicadas en Camilo Ponce Enríquez, Azuay, en las dos primeras instituciones, el proceso de aplicación se realizó en un solo día, mientras que en la Unidad Educativa “Ponce Enríquez” se extendió por aproximadamente un mes, debido a la disponibilidad de los estudiantes, quienes se encontraban en prácticas académicas.

Antes de la aplicación del test SENESCYT y del chatbot, se solicitó autorización a las autoridades de cada institución educativa mediante oficios oficiales, explicando el objetivo del estudio, la voluntariedad de la participación y la seguridad de la información recopilada. La colaboración de los alumnos se realizó respetando estas directrices y bajo supervisión de los investigadores.

### 2.1 Diseño metodológico

Se adopta un diseño cuasiexperimental con grupo único, complementado con un grupo control, lo cual permite fortalecer el análisis comparativo bajo condiciones de aplicación similares. La muestra total se compone de 351 estudiantes del tercer año de bachillerato de instituciones educativas en Ecuador, que realizaron el test SENESCYT y luego interactuaron con el Chatbot.

El muestreo fue no probabilístico por conveniencia, acorde con la disponibilidad institucional, aunque se reconoce que una muestra más representativa podría mejorar la generalización de los resultados.

Cada estudiante del grupo experimental participó en tres etapas consecutivas:

- Primera Fase:** Realización del test vocacional de la SENESCYT.
- Segunda Fase:** Interacción con el chatbot potenciado con inteligencia artificial.
- Tercera Fase:** Aplicación de un test de percepción con escala de Likert.

La encuesta tuvo como meta evaluar los siguientes puntos: claridad de las preguntas, facilidad de uso del chatbot, relevancia de las recomendaciones recibidas, disposición a recomendar la



herramienta a otros estudiantes y nivel de satisfacción sobre la herramienta.

Igualmente, se documentó observaciones directas durante la interacción con el chatbot para enriquecer el análisis cualitativo.

### 2.1.1 Validación previa del chatbot

Antes de la aplicación masiva, se realizaron pruebas internas del sistema por parte de los investigadores para verificar el funcionamiento del chatbot, asegurando la correcta generación de preguntas, la coherencia de las recomendaciones y el adecuado desempeño de la interacción, esta revisión permitió detectar y corregir errores técnicos y optimizar la usabilidad antes de la aplicación con los estudiantes.

### 2.2.2 Condiciones de aplicación

Los instrumentos fueron aplicados en aulas equipadas con computadoras de escritorio provistas por cada institución y se garantizó una conexión a internet estable durante la interacción con el chatbot y el test digital SENESCYT. Los estudiantes realizaron las pruebas de manera individual, bajo supervisión de los investigadores, en un ambiente controlado y sin interrupciones externas, con el fin de mantener la validez ecológica del estudio, cada sesión tuvo una duración aproximada de 40 a 60 minutos, y el número de estudiantes por aula no superó los 25 para asegurar un seguimiento adecuado durante la interacción.

## 2.2 Técnica de análisis

### Análisis cuantitativo

Con la finalidad de examinar el parecido entre las recomendaciones brindadas por ambos instrumentos, se elaboraron cálculos de frecuencia absoluta, porcentajes y medidas de tendencia central, para ello se realizó: Una prueba de hipótesis para lograr identificar diferencias significativas entre ambos grupos y una prueba de Equivalencia TOST (Two One-Sided Tests) que permitirá identificar si las recomendaciones de ambas herramientas son estadísticamente similares.

### Análisis cualitativo

Se lo realizó bajo un proceso de codificación abierta, seguido de una categorización temática de los resultados obtenidos durante las encuestas y observación directa, las dimensiones evaluadas incluyeron empatía, claridad en la comunicación, percepción de utilidad y facilidad de interacción. Por último, el análisis se lo realizó de manera colaborativa por los investigadores, quienes llevaron a cabo revisiones iterativas para poder asegurar la coherencia en la interpretación y minimizar posibles sesgos.

**Tabla 1.** Variables de estudio.

Fuente: Los autores

| Tipo          | Variable  | Instrumento  |
|---------------|---|--|
| Independiente | Tipo de herramienta de orientación                      | Chatbot vocacional<br>Test SENESCYT  |
| Dependiente   | Recomendación vocacional obtenida                       | Resultado del sistema / test   |
| Dependiente   | Nivel de satisfacción del estudiante                    | Encuesta con escala de Likert (5 niveles desde “Totalmente en desacuerdo” hasta “Totalmente de acuerdo”) |
| Dependiente   | Tiempo empleado en la interacción y preguntas adaptadas | Métricas internas del chatbot  |
| Cualitativa   | Percepción sobre la empatía e interacción               | Observación directa y comentarios abiertos   |

### VARIABLES DE ESTUDIO

### 2.3 Secuencia de procedimiento

El procedimiento se organizó basándose en un orden dispuesto:

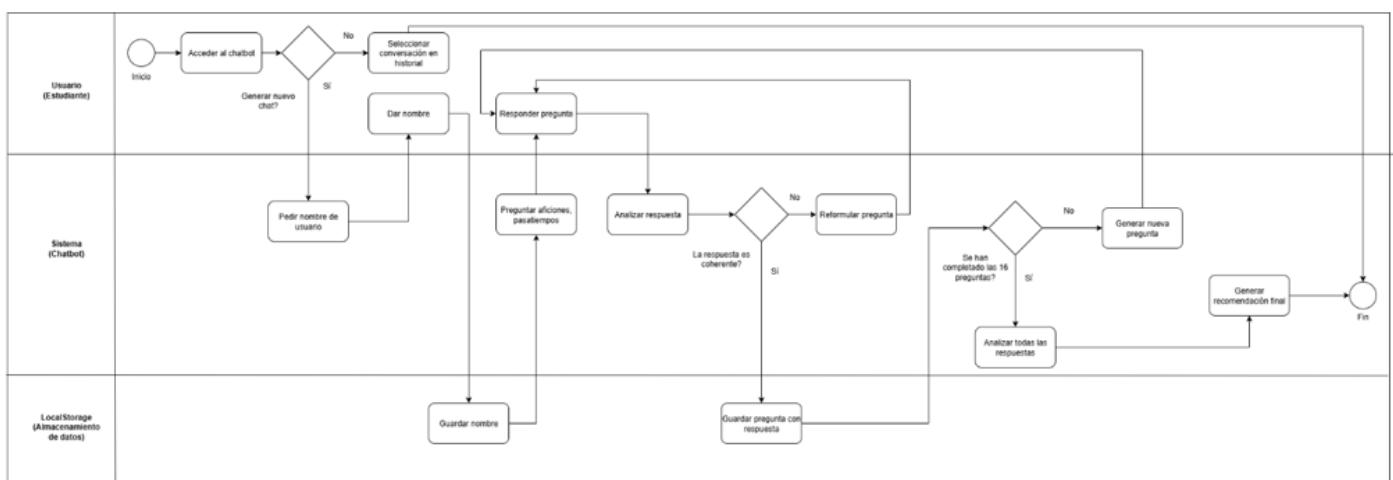
- Aplicación del test SENESCYT (a todos los estudiantes).
- Interacción con el chatbot (solo grupo experimental).
- Encuesta de percepción.
- Observación directa sobre la interacción.
- Recolección de datos en Google Sheets.
- Análisis comparativo.

Cada una de las reuniones fue supervisada por ambos investigadores y gestionada por paralelo, con el propósito de poder mostrar de manera precisa cómo los estudiantes interactuaron con el chatbot y cómo este generó cada una de sus recomendaciones, a continuación, el diagrama funcional que enmarca este proceso se presenta en la Figura 1.

### 2.4 Estructura del sistema

El Chatbot se diseñó con una arquitectura modular y escalable para futuras actualizaciones:

- **Frontend (Angular):** Interfaz del usuario con almacenamiento temporal en localStorage.
- **Backend (Node.js):** Lógica del servidor y conexión con APIs externas.

**Figura 1.** Diagrama funcional del proceso de orientación vocacional que desarrolla el chatbot.

Fuente: Los autores.

- **Motor IA (DeepSeek):** Análisis semántico y generación de recomendaciones con ingeniería de prompts.
- **Base de datos de carreras:** Vinculación con la oferta académica oficial de la UTMACH.
- **Google Sheets API:** Almacenamiento automático de los resultados.
- **Funciones adicionales:** Guardado de la conversación (PDF/Copia en portapapeles).

#### 2.4.1 Seguridad de datos

Para garantizar una gestión adecuada y la protección de la información manejada por el sistema, se implementaron validaciones tanto en el backend como en el frontend, a pesar de que la aplicación no procesa ni gestiona información personal sensible.

#### • Validaciones en el Backend

##### Prevención de duplicados:

En las rutas de la API (/api/chat, /api/opinion y /api/guardar resultado), cada solicitud se procesa de manera controlada, evitando almacenar o reenviar datos duplicados.

##### Manejo seguro de errores

Se utiliza try/catch en cada endpoint para capturar excepciones y devolver mensajes genéricos al cliente. Esto evita exponer detalles internos del servidor (como trazas de error completas) que podrían representar un riesgo de seguridad.

##### Control de acceso mediante CORS:

Se habilitó cors() en Express para definir qué orígenes tienen permitido interactuar con la API. Esta medida previene solicitudes no autorizadas desde otros dominios.

#### • Validaciones en el Frontend

##### Filtrado de conversaciones antes de guardar:

El servicio ChatStorageService realiza un filtrado previo, cada vez que se guarda una conversación en localStorage:

- Lee todas las conversaciones existentes desde el navegador.
- Filtra la que pertenece al mismo usuario para evitar duplicación.
- Limpia los mensajes eliminando campos como html, reduciendo el riesgo de inyección de código.
- Guarda solo la versión actualizada y validada.

##### Normalización de datos:

Se comparan los nombres de usuario de manera insensible a mayúsculas/minúsculas y espacios en blanco, evitando duplicaciones por variaciones de escritura (por ejemplo, “Juan” y “juan”).

#### 2.5 Desglose de sprints y cronograma

Se utilizaron 6 sprints de 2 semanas cada uno, conforme al marco de trabajo ágil Scrum. Cada sprint tuvo entregables definidos. Se incorporaron indicadores de evaluación por sprint, como porcentaje de tareas completadas, tiempo de respuesta del sistema y satisfacción del usuario.

##### Sprint 1 – Diseño estructural y flujo conversacional

**Duración:** 2 semanas

**Objetivo principal:** Construcción de la interfaz inicial y la lógica básica de conversación.

**Fundamento:**

Se evaluaron las habilidades en comunicación, matemáticas, competencias digitales y socioemocionales de los estudiantes según la rama común del bachillerato, con el propósito de elaborar preguntas abiertas alineadas a las directrices del currículo priorizado establecido por el Ministerio de Educación del Ecuador.



**Tareas:**

- Revisión del currículo priorizado centrado en las competencias clave de los estudiantes.
- Diseño de preguntas abiertas (sin IA).
- Desarrollo de la interfaz en Angular.
- Simulación del flujo conversacional con preguntas estáticas.
- Pruebas de navegación y usabilidad.

**Entregables:**

- Interfaz funcional inicial.
- Flujo conversacional simulado con preguntas estáticas.
- Primeras pruebas de interacción con el sistema.

**Sprint 2 – Integración de la inteligencia artificial (DeepSeek) en el sistema de chat.****Duración:** 2 semanas

**Objetivo principal:** Realizar la integración de DeepSeek para evaluar las respuestas de los estudiantes y generar preguntas de manera adaptativa según sus intereses.

**Fundamento:**

Se escogió Deepseek por su capacidad de interpretar respuestas abiertas, formular preguntas personalizadas y entregar recomendaciones alineadas con el perfil de cada estudiante.

**Tareas:**

- Construcción de prompts dinámicos.
- Conexión con DeepSeek.
- Sustitución de preguntas estáticas por generación automatizada.
- Procesamiento de respuestas y generación de recomendaciones.
- Validaciones funcionales y de seguridad.

**Entregables:**

- Flujo de interacción completamente adaptable según las respuestas de los alumnos.
- Recomendaciones vocacionales personalizadas según el perfil de cada estudiante.

**Sprint 3 – Desarrollo de un flujo de interacción adaptativo y eficiente****Duración:** 2 semanas

**Objetivo principal:** Mejorar la experiencia de interacción,

asegurando que las respuestas sean coherentes y transmitan empatía hacia los estudiantes.

**Fundamento:**

Se procuró que el sistema ofreciera una interacción similar a una conversación humana, identificando ambigüedades, reformulando preguntas cuando era necesario y ajustando el lenguaje a cada uno de los estudiantes mediante algoritmos de análisis semántico del lenguaje.

**Tareas:**

- Verificación de respuestas que puedan ser ambiguas o incompletas.
- Reformulación automática de preguntas según la interacción del estudiante.
- Optimización de la empatía y claridad durante la interacción conversacional.
- Ajustes en la interfaz para mejorar la accesibilidad y comprensión del usuario.
- Realización de pruebas de interacción con estudiantes para evaluar la efectividad del sistema.
- Implementar almacenamiento local (localStorage) para mantener el estado del chat.
- Desarrollo de un sistema que permite guardar y recuperar el historial de conversaciones de manera local.
- Validar la usabilidad y accesibilidad de estas funciones con usuarios.

**Entregables:**

- Flujo conversacional adaptativo y empático.
- Mejora en naturalidad y conexión con el usuario.
- Persistencia y recuperación de estado de chat funcionando.
- Funciones de exportación y copia implementadas y probadas.
- Avances en la experiencia de usuario y accesibilidad confirmadas.

**Sprint 4 – Registro de resultados finales en Google Sheets****Duración:** 2 semanas

**Objetivo principal:** Almacenar externamente los resultados vocacionales sugeridos por la IA.

**Fundamento:**

Para facilitar el seguimiento institucional y el análisis estadístico,

se integró Google Sheets como repositorio externo para registrar únicamente el nombre del estudiante y la carrera sugerida.

#### Tareas:

- Programación del backend para envío automático de datos.
- Registro del nombre del estudiante y la recomendación al completar el flujo.
- Validaciones para evitar duplicidad.
- Pruebas de integridad y consistencia de datos.

#### Entregables:

- Registro automático de resultados finales en Google Sheets.
- Documento centralizado para uso institucional.

### Sprint 5 – Implementación del test tradicional SENESCYT (comparativo)

Duración: 2 semanas

**Objetivo principal:** Construir una versión web del test tradicional SENESCYT para comparación directa con el Chatbot.

#### Fundamento:

Con el propósito de validar la eficacia del Chatbot, se implementó una versión digital del test vocacional tradicional aplicado por la SENESCYT. Esto permitió recoger resultados bajo condiciones similares a las del Chatbot, facilitando una comparación justa y ordenada.

#### Tareas:

- Digitalización completa del test SENESCYT (estructura de preguntas cerradas).
- Desarrollo de una interfaz web accesible y funcional para los estudiantes.
- Registro automatizado de resultados (carrera sugerida) en Google Sheets.
- Asignación de identificadores únicos a cada estudiante para vincular los resultados de ambos métodos.
- Preparación de los registros para análisis comparativo y validación cruzada.

#### Entregables:

- Plataforma funcional y usable del test vocacional SENESCYT en formato web.
- Resultados organizados en Google Sheets, diferenciados de los del chatbot.
- Registro estructurado que permite comparar recomendaciones de ambos instrumentos con base en un mismo identificador.

### Sprint 6 – Validación final y análisis de impacto

Duración: 2 semanas planificadas, extendidas a 4 semanas debido a la disponibilidad de los estudiantes.

**Objetivo principal:** Evaluar la efectividad del chatbot frente al test SENESCYT con estudiantes reales.

#### Fundamento:

Se desarrolló una prueba piloto con 351 estudiantes de tercer año de bachillerato, utilizando ambos instrumentos. Los resultados permitieron comparar coincidencias, divergencias, y percepciones sobre utilidad, comprensión y preferencia.

#### Tareas:

- Aplicación paralela de ambos instrumentos.
- Recopilación de resultados cuantitativos y observaciones cualitativas.
- Análisis de correlación y nivel de coincidencia.
- Ajustes finales al chatbot.
- Redacción de informe de validación.

#### Entregables:

- Sistema probado en condiciones reales.
- Informe de validación y recomendaciones para implementación futura.

## 2.6 Herramientas usadas durante el desarrollo

- **Visual Studio Code:** Entorno principal de desarrollo para frontend y backend, con extensiones específicas para Angular, TypeScript y control de versiones mediante GIT.
- **Angular:** Framework utilizado para construir una interfaz modular, responsive e interactiva que facilita una experiencia de usuario clara e intuitiva.
- **Node.js:** Entorno de ejecución backend para gestionar la lógica del servidor, conexión con APIs externas y automatización de tareas.
- **API DeepSeek:** Interfaz de inteligencia artificial utilizado para interpretar respuestas y generar recomendaciones vocacionales para cada uno de los usuarios.
- **LocalStorage (del navegador):** Herramienta utilizada para poder almacenar temporalmente el chat de cada estudiante.
- **Google Sheets API:** Servicio agregado para guardar automáticamente cada uno de los resultados vocacionales de los estudiantes (tanto chatbot como el test SENESCYT), facilitando su posterior análisis.
- **Vercel:** Plataforma utilizada para lograr alojar el backend de la aplicación el cual incluye las funciones con DeepSeek y Google Sheets.
- **Scrum Board digital:** Herramienta utilizada para poder planificar, distribuir y dar seguimiento a las tareas de cada sprint dentro del equipo de desarrollo.



## 2.7 Entregables por Sprint

La Tabla 2 resume cada uno de los seis Sprints que se desarrollaron siguiendo la metodología ágil Scrum, cada una de las iteraciones contó con entregables específicos, medibles y verificables, lo que permitió poder avanzar de manera incremental en la construcción del sistema. El estado “Completado” permite reflejar que las actividades planificadas en cada sprint se lograron finalizar con éxito, permitiendo que el chatbot vocacional pueda evolucionar desde un prototipo inicial hasta una versión probada en condiciones reales. Esta gestión facilitó la trazabilidad del proyecto, el seguimiento de la calidad en cada fase y la detección temprana de oportunidades de mejora, asegurando que el producto final cumpliera con los objetivos planteados en la investigación.

**Tabla 2.** Entregables por Sprint del Proyecto.

Fuente: Los autores.

| Sprint   | Entregable Principal                               | Estado     | Descripción Breve   |
|----------|--|------------|---|
| Sprint 1 | Interfaz inicial + flujo básico                    | Completado | Interfaz en Angular con navegación conversacional básica y almacenamiento local.                  |
| Sprint 2 | Integración con IA DeepSeek                        | Completado | IA procesa respuestas abiertas y entrega recomendaciones personalizadas.                          |
| Sprint 3 | Conversación empática e inteligente                | Completado | Flujo dinámico con validación semántica y reformulación de preguntas.                             |
| Sprint 4 | Registro automático de resultados en Google Sheets | Completado | Resultados de los estudiantes almacenados automáticamente para su análisis institucional.         |
| Sprint 5 | Plataforma web del test SENESCYT                   | Completado | Versión digital del test tradicional vinculada a Google Sheets para el seguimiento de resultados. |
| Sprint 6 | Validación y análisis comparativo                  | Completado | Aplicación ambos instrumentos a estudiantes y recopilación de percepciones sobre su utilidad.     |

## 3. Resultados y Discusión

Al concluir con la fase de implementación se procedió a aplicar tanto el instrumento oficial de la SENESCYT como la herramienta de chatbot a un total de 351 estudiantes del tercer año de bachillerato, esto permitió realizar un análisis cuantitativo mediante técnicas estadísticas, entre ellas la prueba de hipótesis por proporciones y por último el Test de Equivalencia (TOST). Asimismo, se pudo efectuar un análisis cualitativo comparativo que permitió evaluar los resultados obtenidos, la percepción de los estudiantes y la eficacia de las recomendaciones emitidas.

### 3.1 Prueba de hipótesis

#### 3.1.1 Definición del problema

- Población: Estudiantes de tercero de bachillerato.
- Muestra: n = 351
- Resultados coincidentes (chatbot vs. SENESCYT): x=281.
- Proporción muestral:

$$\hat{p} = \frac{281}{351} = 0.8006 \approx 80.1\% \quad (1)$$

#### 3.1.2 Hipótesis estadísticas

La hipótesis se enfoca en determinar si el chatbot tiene al menos la misma efectividad que el test SENESCYT en términos de coincidencia de recomendaciones, podemos asumir como valor de referencia que el test SENESCYT tiene una efectividad del 80% (ya que lo estamos usando como estándar). Entonces:

##### • Hipótesis nula ( $H_0$ ):

$$p \leq 0.80$$

El chatbot no es igual ni mejor que el test SENESCYT.

##### • Hipótesis alternativa ( $H_1$ ):

$$p \geq 0.80$$

El chatbot tiene un rendimiento igual o mejor que el test SENESCYT.

Esto corresponde a una prueba de hipótesis unilateral (cola derecha).

#### 3.1.3 Estadístico de prueba

Se usa la prueba z para una proporción:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

donde:

- $\hat{p} = 0.8006$
- $p_0 = 0.80$
- $n = 351$

Cálculo:

1. Error estándar (SE) para la proporción:

$$SE = \sqrt{\frac{0.8(0.2)}{351}} = \sqrt{\frac{0.16}{351}} \approx 0.02136 \quad (3)$$

2. Cálculo del estadístico z:

$$z = \frac{0.8006 - 0.80}{0.02136} = \frac{0.0006}{0.02136} \approx 0.028 \quad (4)$$

**Valor p:** Con  $z = 0.028$  el valor p (cola derecha) es aproximadamente:

p-valor=0.489≈49%

**Decisión:** Si usamos un nivel de significancia clásico ( $\alpha=0.05$ ):

Como p-valor=0.489> 0.05 no se rechaza  $H_0$ .

### 3.1.4 Interpretación

Con los datos actuales, no se observa la suficiente evidencia, desde el punto de vista estadístico, que permita afirmar que el chatbot supere el test de la SENESCYT.

Lo que sí se puede afirmar es que el chatbot alcanza prácticamente la misma precisión (80.1%) que el test (80%), es decir, se comporta de manera equivalente en este estudio piloto.

## 3.2 Prueba de equivalencia (TOST)

### Datos

- Muestra:  $n = 351$ .
- Coincidencias:  $x=281$ .
- Proporción observada:

$$\hat{p} = \frac{281}{351} = 0.8006 \quad (5)$$

• Valor de referencia:  $p_0 = 0.80$ .

### 3.2.1 Hipótesis de equivalencia

Debemos definir un margen de equivalencia ( $\delta$ ), que es la tolerancia que se acepta como diferencia “no significativa” entre el chatbot y el test.

En estudios de educación o ciencias sociales suele usarse  $\pm 5\% (\delta=0.05)$ .

Entonces:

- $H_0: p \leq 0.75$  o  $p \geq 0.85$  (el chatbot no es equivalente al test).
- $H_1: 0.75 < p < 0.85$  (el chatbot es equivalente al test dentro del margen del 5%).

### 3.2.2 Intervalo de confianza

Para la prueba TOST, se construye un IC al 90% (porque se hacen dos pruebas unilaterales).

El error estándar es:

$$SE = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.8006 \cdot 0.1994}{351}} \approx 0.0213 \quad (6)$$

El margen para un 90% IC con  $z_{0.95}=1.645$ :

$$ME = 1.645 \times 0.0213 \approx 0.035 \quad (7)$$

Por lo tanto, el IC (90%) es:

$$IC_{90\%} = (0.8006 - 0.035, 0.8006 + 0.035) = (0.7656, 0.8356) \quad (8)$$

### 3.2.3 Decisión

- El IC90% del chatbot está entre 76.6% y 83.6%.
- El margen de equivalencia era [0.75, 0.85].
- Como el IC está contenido dentro del margen, rechazamos  $H_0$  y aceptamos  $H_1$ .

### 3.2.4 Interpretación

Con un nivel de confianza del 90%, se demuestra que la precisión del chatbot es estadísticamente equivalente a la del test vocacional de la SENESCYT dentro de un margen de  $\pm 5\%$ .

En otras palabras, el chatbot no es inferior al test, y se puede considerar una herramienta válida y alternativa para la orientación vocacional.

## 3.3 Comparación entre chatbot y test SENESCYT con base en la encuesta con escala Likert.

Los resultados obtenidos evidencian que, si bien ambos instrumentos coincidieron en un alto nivel respecto a las recomendaciones vocacionales, la percepción de los estudiantes fue ampliamente favorable hacia el Chatbot en comparación con el test SENESCYT.

En la gráfica se observa que las barras correspondientes al Chatbot (color verde) superan consistentemente al test SENESCYT (color azul) en todos los criterios evaluados: confianza, modernidad, claridad de preguntas, experiencia de usuario, satisfacción general, empatía y disposición a recomendar la herramienta.

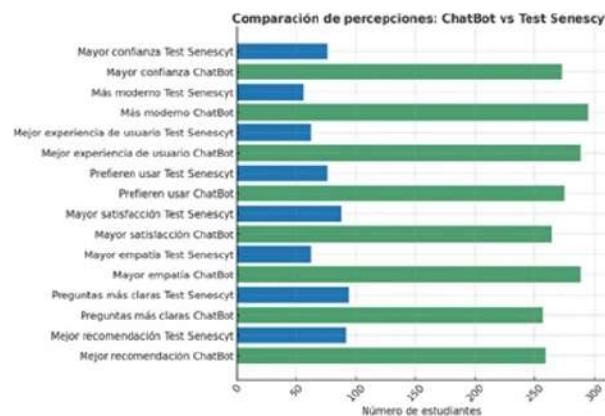
Este resultado refleja que los estudiantes no solo consideran al Chatbot tan preciso como el test tradicional, sino también más dinámico, empático y comprensible, lo cual coincide con los



**Tabla 3.** Interpretación de resultados

Fuente: Los autores

| Característica                              | Chatbot<br>(Proporción observada) | Test SENESCYT<br>(Proporción de Referencia) | Margen de<br>equivalencia ( $\delta$ ) | Intervalo de Confianza<br>del 90% (IC) | Decisión TOST              |
|---|-----------------------------------|---|--|--|----------------------------|
| Proporción de Coincidencia<br>(Efectividad) | 80,1%<br>( $p \approx 0.801$ )    | 80.0%<br>( $p_0 = 0.80$ )                   | $\pm 5\%$ (0.05)                       | [76.6%; 83.6%]                         | Equivalencia<br>demostrada |



**Figura 2.** Comparación de percepciones entre las dos herramientas.

Fuente: Los autores

hallazgos de análisis globales que muestran un efecto positivo de los chatbots sobre el aprendizaje y la experiencia estudiantil (Debets et al., 2025; Laun & Wolff, 2025; Yu & Zhao, 2025). Esto sugiere que la interacción conversacional propia de los sistemas de IA facilita la identificación personal con las recomendaciones, aumentando la confianza en el proceso.

No obstante, se debe enfatizar que el uso de sistemas automatizados implica considerar principios éticos vinculados con la equidad, transparencia y resguardo de la información del estudiante (Al-Zahrani, 2025; Holmes et al., 2022), a fin de garantizar un uso responsable y confiable en contextos educativos.

#### 3.4 Interfaz del sistema

La interfaz del chatbot fue diseñada para proporcionar una experiencia conversacional fluida, amigable y clara, este tipo de personalización conversacional responde a principios de adaptabilidad descritos en estudios recientes sobre sistemas educativos impulsados por IA, que enfatizan la necesidad de ajustar la experiencia según las respuestas del usuario. Las ilustraciones a continuación muestran momentos clave del flujo conversacional.



**Figura 3.** Inicio de la conversación con el chatbot.

Fuente: Los autores.



**Figura 4.** Recomendación vocacional personalizada generada por IA.

Fuente: Los autores.

#### 3.5 Grado de satisfacción de los estudiantes

Para conocer la percepción de los participantes, se aplicó una encuesta basada en una escala tipo Likert. Los resultados



**Figura 5.** Grado de satisfacción general con el chatbot.

Fuente: Los autores

reflejaron una valoración muy positiva hacia el sistema. Los estudiantes evalúan aspectos como la claridad de las indicaciones, la facilidad de interacción, la pertinencia de las recomendaciones y su disposición a sugerir el uso del chatbot a otros estudiantes. La distribución de las respuestas relacionadas con la experiencia de uso se presenta en el gráfico circular correspondiente.

### 3.6 Comentarios y observaciones cualitativas

Durante la aplicación del instrumento se recopilaron comentarios espontáneos de los estudiantes, los cuales permitieron complementar el análisis cuantitativo con apreciaciones de carácter cualitativo. Entre ellos se destacaron expresiones como:

- “Me sentí más comprendido por el chatbot que con el test tradicional”.
- “Las preguntas del chatbot eran más abiertas, me hicieron pensar en lo que realmente me gusta”.
- “La interfaz es sencilla y agradable, y resultó fácil responder”.

Estas opiniones refuerzan los resultados obtenidos en el análisis estadístico, al evidenciar una mayor conexión emocional de los participantes con el sistema. Investigaciones previas en el campo de la orientación vocacional han reportado hallazgos afines, señalando que los chatbots favorecen procesos de aprendizaje más personalizados y eficaces (Hoang et al., 2023).

De manera concordante, otros estudios realizados en entornos de aula que incorporan chatbots para promover habilidades reflexivas reportaron altos niveles de aceptación y satisfacción estudiantil. Tales percepciones se alinean con las tendencias actuales en el uso de la inteligencia artificial en educación, en las que se valora la interacción personalizada como complemento a las metodologías tradicionales (Martínez-Olmo & Catalán, 2024).

### 3.7 Limitaciones del Estudio

- El uso de un muestreo no probabilístico por conveniencia restringe la generalización de los hallazgos al total de la población estudiantil.
- La comparación entre preguntas abiertas (chatbot) y cerradas (test SENESCYT) podría afectar la validez del análisis comparativo.
- La arquitectura del sistema no incluye aprendizaje automático continuo, lo que impide mejorar progresivamente las recomendaciones según la interacción real de los estudiantes.
- El análisis cualitativo, aunque realizado por dos investigadores, puede contener sesgos interpretativos; se contó con revisión externa de tutores para reforzar la consistencia de los resultados.

### 3.8 Propuestas para mitigar limitaciones

- Ampliar la muestra e incluir instituciones de diferentes regiones y tipos (públicas y privadas).

- Implementar un modelo de aprendizaje continuo en el motor IA para mejorar la precisión de las recomendaciones.
- Incluir revisiones externas por tutores o expertos, con el fin de garantizar la validez interpretativa del análisis cualitativo y la consistencia de los resultados.

### 3.9 Síntesis de resultados y discusión

En general, los hallazgos permiten destacar lo siguiente:

- 1. Equivalencia en efectividad:** El chatbot logró alcanzar la misma precisión que el test SENESCYT en cuanto las recomendaciones vocacionales.
- 2. Superación en percepción:** Los estudiantes calificaron al chatbot de una manera mucho más alta durante el test de escala tipo Likert, lo cual permitió mostrar una mayor aceptación y confianza sobre el sistema.
- 3. Valor agregado en la experiencia:** La interacción conversacional y la percepción de empatía de los estudiantes generaron una sensación de cercanía que no se experimenta con la prueba tradicional.
- 4. Validez como alternativa:** El chatbot demostró ser efectivo y bien recibido por los estudiantes, ofreciéndoles un complemento más innovador que los métodos tradicionales de orientación vocacional.
- 5. Carácter comparativo y contraste con estudios internacionales:** El rendimiento del Chatbot refuerza su carácter comparativo al alinearse con la tendencia global. La demostración de la equivalencia en precisión valida el uso de sistemas basados en IA en la orientación, como se ha observado en estudios de Corea del Sur y Europa, mientras que la alta satisfacción del usuario refleja la preferencia por la disponibilidad y la inmediatez de los asistentes virtuales en la educación superior observadas en universidades de Estados Unidos y Reino Unido. Este estudio confirma que el modelo de Chatbot es una solución escalable y costo-efectivo que moderniza la orientación vocacional conforme a los estándares de innovación educativa internacional.

En conjunto, estos resultados respaldan la viabilidad del uso de chatbots educativos en procesos de orientación y destacan la necesidad de investigaciones futuras que incluyan muestras más amplias y evalúen los efectos a largo plazo.

### 4. Conclusiones

Los resultados de esta investigación concluyen que el Chatbot desarrollado es una herramienta innovadora y confiable para la orientación vocacional. El análisis estadístico, validado con un nivel de confianza del 90%, demostró que el desempeño del Chatbot es estadísticamente comparable al de la prueba tradicional de la SENESCYT, lo que prueba la confiabilidad diagnóstica del sistema de IA.



Además de su precisión, la herramienta evidencia un alto valor agregado: los estudiantes expresan una clara preferencia por su uso, resaltando su claridad, la percepción de empatía y la facilidad de interacción. Este doble hallazgo subraya el potencial del Chatbot como una alternativa innovadora y accesible para el proceso de elección de carrera.

Más allá de los resultados cuantitativos, este estudio resalta el impacto educativo y social que podría generar la adopción de soluciones digitales interactivas. A diferencia de los métodos tradicionales, el Chatbot no solo entrega resultados similares, sino que su implementación es especialmente relevante en instituciones públicas o rurales con escasez de orientadores vocacionales. En estos entornos, la herramienta ofrece una solución de bajo costo y alta escalabilidad que garantiza el acceso a una orientación académica de calidad y personalizada para todos los estudiantes. Este enfoque representa un paso importante hacia la modernización de la orientación académica en contextos de vulnerabilidad o limitación de recursos.

Finalmente, los resultados de este proyecto abren la puerta a líneas futuras de investigación enfocadas en la sostenibilidad tecnológica del sistema.

Específicamente, se sugiere explorar la implementación de aprendizaje continuo y la retroalimentación adaptativa del Chatbot. Estas mejoras permitirían que la herramienta perfeccione progresivamente la precisión de sus recomendaciones con cada interacción, asegurando su relevancia a largo plazo y motivando a instituciones como la Universidad Técnica de Machala a considerar la adopción de herramientas tecnológicas que faciliten una orientación vocacional más precisa, relevante y acorde con los retos educativos actuales.

### Agradecimientos

Quiero dar las gracias a la Universidad Técnica de Machala por todo el apoyo y los recursos que me dieron para poder hacer esta investigación. También extiendo mi gratitud a las autoridades de la universidad; su visión es la que hace posible que la innovación y el crecimiento académico sigan adelante. Y, por supuesto, un agradecimiento muy especial a todos los estudiantes y sus instituciones educativas que participaron en este estudio. Su colaboración fue clave y, sin su ayuda, no habría sido posible recolectar los datos que respaldan los resultados de este trabajo.

### Contribución de los autores

**Fernando Ericsson Montoya Naguas:** Conceptualización, Metodología, Software, Investigación, Curación de datos, Análisis formal, Redacción - borrador original del artículo. **Carlos Elian Lalangui Flores:** Conceptualización, Metodología, Software,

Investigación, Curación de datos, Análisis formal, Redacción - revisión y edición del artículo, Validación. **Fausto Fabián Redrován Castillo:** Supervisión, Redacción - revisión y edición del artículo. **Freddy Aníbal Jumbo Castillo:** Supervisión, Redacción - revisión y edición del artículo.

### Conflictos de interés

Los autores no tienen conflictos de interés respecto al presente estudio.

### Referencias bibliográficas

- Al-Zahrani, A. M. (2025). Exploring the Impact of Artificial Intelligence Chatbots on Human Connection and Emotional Support Among Higher Education Students. *Sage Open*, 15. <https://doi.org/10.1177/21582440251340615>
- Caldarini, G., Jaf, S., & McGarry, K. (2022). A Literature Survey of Recent Advances in Chatbots. *Information (Switzerland)*, 13. <https://doi.org/10.3390/info13010041>
- Davar, N. F., Dewan, M. A. A., & Zhang, X. (2025). AI Chatbots in Education: Challenges and Opportunities. *Information (Switzerland)*, 16(3). <https://doi.org/10.3390/info16030235>
- Debets, T., Banihashem, S. K., Joosten-Ten Brinke, D., Vos, T. E. J., Maillette de Buy Wenniger, G., & Camp, G. (2025). Chatbots in education: A systematic review of objectives, underlying technology and theory, evaluation criteria, and impacts. *Computers and Education*. <https://doi.org/10.1016/j.compedu.2025.105323>
- Erazo, X., & Rosero, E. (2021). Orientación vocacional y su influencia en la deserción universitaria. *Horizontes. Revista de Investigación en Ciencias de la Educación*, 5, 591-606. <https://doi.org/10.33996/revistahorizontes.v5i18.198>
- Haristiani, N., Dewanty, V. L., & Rifai, M. M. (2022). Autonomous Learning Through Chatbot-based Application Utilization to Enhance Basic Japanese Competence of Vocational High School Students. *Journal of Technical Education and Training*, 14(2 Special Issue), 143-155. <https://doi.org/10.30880/jtet.2022.14.02.013>
- Hoang, N. T., Ngoc Han, D., & Le, D. H. (2023). Exploring Chatbot AI in improving vocational students' English pronunciation. *AsiaCALL Online Journal*, 14(2), 140-155. <https://doi.org/10.54855/acoj.231429>

- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in Education: Towards a Community-Wide Framework. *International Journal of Artificial Intelligence in Education*, 32(3), 504-526. <https://doi.org/10.1007/s40593-021-00239-1>
- Jaramillo, S. Y. A. (2024). Orientación Vocacional y Profesional: La Inteligencia Artificial y su impacto en la Educación. *Revista Scientific*, 9(34), 285-300. <https://doi.org/10.29394/scientific.issn.2542-2987.2024.9.34.13.285-300>
- Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. En International Journal of Educational Technology in Higher Education (Vol. 20, Número 1). Springer Science and Business Media Deutschland GmbH. <https://doi.org/10.1186/s41239-023-00426-1>
- Laun, M., & Wolff, F. (2025). Chatbots in education: Hype or help? A meta-analysis. *Learning and Individual Differences*, 119. <https://doi.org/10.1016/j.lindif.2025.102646>
- Lee, Y. F., Hwang, G. J., & Chen, P. Y. (2025). Technology-based interactive guidance to promote learning performance and self-regulation: a chatbot-assisted self-regulated learning approach. *Educational Technology Research and Development*. <https://doi.org/10.1007/s11423-025-10478-x>
- Lucana, Y., & Roldan, W. (2023). Chatbot basado en inteligencia artificial para la educación escolar. Horizontes. *Revista de Investigación en Ciencias de la Educación*, 7, 1580-1592. <https://doi.org/10.33996/revistahorizontes.v7i29.614>
- Martínez-Olmo, F., & Catalán, F. G. (2024). Systematic Review of Trends in the Application of Artificial Intelligence to the Field of Academic Writing in the Social Sciences. *Digital Education Review*, 45, 37-42. <https://doi.org/10.1344/der.2024.45.37-42>
- Mayon, J., & Ruiz, M. (2024). Desarrollo de una aplicación de orientación vocacional para bachilleres mediante test utilizando inteligencia artificial para la toma de decisiones. [Universidad Técnica de Machala]. [https://repositorio.utmachala.edu.ec/bitstream/48000/23575/1/Trabajo\\_Titulacion\\_3439.pdf](https://repositorio.utmachala.edu.ec/bitstream/48000/23575/1/Trabajo_Titulacion_3439.pdf)
- Ministerio de Educación del Ecuador. (2025). Currículo Priorizado con Énfasis en Competencias Comunicacionales, Matemáticas, Digitales y Socioemocionales. Ministerio de Educación del Ecuador. [www.educacion.gob.ec](http://www.educacion.gob.ec)
- Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2. <https://doi.org/10.1016/j.caai.2021.100033>
- Phuttwawong, M., & Chatwattana, P. (2025). The Educational Guidance Platform via Artificial Intelligence Chatbot to Promote Vocational Aptitude for Vocational Students. *Higher Education Studies*, 15(1), 128-135. <https://doi.org/10.5539/hes.v15n1p128>
- Torres-Sevillano, D. M., Vilcacundo-Gallardo, F. R., & García-Hevia, S. (2025). Orientación vocacional en la elección de carreras técnicas para estudiantes de 10mo grado. *MQRInvestigar*, 9. <https://doi.org/10.56048/mqr20225.9.2.2025.e459>
- Toscu, S. (2024). An Investigation on the Effectiveness of Chatbots in Evaluating Writing Assignments in EFL Contexts. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 295-329. <https://doi.org/10.21764/maeufd.1425384>
- Yu, Y., & Zhao, Z. (2025). Development of vocational education and training chatbot supported by large language model-based multi-agent system. *Vocation, Technology & Education*. <https://doi.org/10.54844/vte.2025.0921>





## Prototipo de Sistema de Monitoreo Inteligente para la detección en Tiempo Real de Objetos y Actividades Sospechosas Utilizando Deep Learning

### Prototype of an Intelligent Monitoring System for Real-Time Detection of Suspicious Objects and Activities Using Deep Learning

#### Autores

\*Lauro Alfonso Erreyes Cuenca

 [lerreyes1@utmachala.edu.ec](mailto:lerreyes1@utmachala.edu.ec)



Nahin Josue Olmedo Chica

 [nolmedo2@utmachala.edu.ec](mailto:nolmedo2@utmachala.edu.ec)



Mariuxi Paola Zea Ordóñez

 [mzea@utmachala.edu.ec](mailto:mzea@utmachala.edu.ec)



Nancy Magaly Loja Mora

 [nmloja@utmachala.edu.ec](mailto:nmloja@utmachala.edu.ec)



Universidad Técnica de Machala,  
Facultad de Ingeniería Civil, Carrera de  
Tecnologías de la Información, Machala,  
El Oro, Ecuador.

\*Autor para correspondencia

#### Resumen

Esta investigación trata sobre la necesidad de optimizar los sistemas de videovigilancia a través del uso de inteligencia artificial para detectar proactivamente amenazas. El propósito primordial es desarrollar un prototipo de monitoreo inteligente que pueda detectar en tiempo real a personas, armas y conductas sospechosas, enfocándose en la eficiencia computacional y la precisión de detección para garantizar su viabilidad en hardware accesible. Se utilizó la metodología CRISP-DM para conseguir esta meta, que dividió el proyecto de manera sistemática en etapas que incluyan la preparación, modelado y valoración de los datos. Un modelo YOLOv8 es el elemento principal del sistema, el cual fue entrenado con un conjunto de datos personalizado que comprende cerca de 8.500 imágenes y se expandió a través de distintos métodos. La robustez del modelo se confirma mediante los resultados cuantitativos, que muestran una puntuación F1-score del 93,99% y una precisión media (mAP) de 50 del 96,97% en las clases especificadas. Finalmente, el modelo fue incorporado en un prototipo funcional de videovigilancia, lo que demostró su utilidad y efectividad operativa en entornos de seguridad urbana y comercial.

**Palabras clave:** inteligencia artificial; deep learning; YOLOv8; videovigilancia.

#### Abstract

This research addresses the need to optimize video surveillance systems through the use of artificial intelligence to proactively detect threats. The primary goal is to develop an intelligent monitoring prototype capable of detecting people, weapons, and suspicious behavior in real time, focusing on computational efficiency and detection accuracy to ensure its feasibility on readily available hardware. The CRISP-DM methodology was used to achieve this goal, systematically dividing the project into stages that included data preparation, modeling, and evaluation. A YOLOv8 model is the core element of the system, trained on a custom dataset of approximately 8,500 images and expanded using various methods. The model's robustness is confirmed by quantitative results, which show an F1 score of 93.99% and a mean accuracy (mAP) of 96.97% in the specified classes. Finally, the model was incorporated into a functional video surveillance prototype, demonstrating its usefulness and operational effectiveness in urban and commercial security environments.

**Keywords:** artificial intelligence; deep learning; YOLOv8; video surveillance.

Enviado: 27/09/2025

Aceptado: 18/11/2025

Publicado: 19/11/2025



## 1. Introduction

Con el rápido crecimiento de la Inteligencia Artificial (IA) en los últimos años, se han generado importantes avances en diversas áreas, como la de videovigilancia inteligente, donde los sistemas tradicionales han quedado atrás. Esto principalmente por la dependencia de la supervisión humana, lo que genera retrasos y errores en situaciones críticas. Un claro ejemplo de cómo afrontar estas debilidades es el desarrollo de sistemas de seguridad capaces de funcionar de manera automatizada y sin supervisión mediante modelos de detección en tiempo real basados en deep learning (Azatbekuly et al., 2024). Estos trabajos resaltan que, a diferencia de los métodos tradicionales que requieren mucha mano de obra para procesarlo, analizarlo y verlo, es posible reducir significativamente la intervención humana y mejorar la capacidad de respuesta ante diversos escenarios.

La videovigilancia apoyada con IA se ha desplegado en escenarios muy variados, no únicamente se limita a ciudades. En el ámbito agrícola, se han creado plataformas que tienen la capacidad de monitorear e identificar plagas en los cultivos automáticamente con un 92% de precisión (Delwar et al., 2025). Un sistema de vigilancia para una mina de carbón en el sector minero se ha puesto en marcha; un estudio reciente demuestra que este tipo de sistemas es capaz de identificar la existencia de individuos en zonas peligrosas y restringidas, con una tasa de precisión impresionante del 99.5%, lo que deja claro su eficacia aun en situaciones operativas extremas (Ni et al., 2024). Al mismo tiempo, se han creado sistemas semejantes en el sector educativo para detectar comportamientos anormales, como conflictos físicos y actos de deshonestidad académica. Estas tecnologías innovadoras han alcanzado un grado de exactitud del 96%, lo que demuestra que pueden proporcionar una respuesta precisa y optimizar las medidas de seguridad en los espacios universitarios (Gawande et al., 2024). Estos ejemplos demuestran que la IA puede ayudar a la seguridad de manera confiable y flexible cuando se utiliza en videovigilancia.

YOLOv8 es un instrumento ideal para sistemas de vigilancia en tiempo real debido a su balance sobresaliente entre precisión y velocidad de procesamiento. Su eficacia se ha comprobado a través de múltiples investigaciones empíricas. Por ejemplo, este modelo se ha empleado para identificar armas de fuego y monitorizar conductas sospechosas en la práctica, logrando un 92.2% de precisión con una Intersección sobre Unión de 0.6 (Schcolnik-Elias et al., 2023). En consecuencia, otros estudios resaltan que YOLOv8 mantiene un rendimiento constante incluso en situaciones adversas. Por ejemplo, se ha confirmado que el mAP@50 alcanzó un 89% y los FPS fueron más de 430 en situaciones con una alta afluencia, lo cual evidencia su

efectividad (Hua et al., 2024). En espacios con una alta densidad de población, también se ha verificado la precisión y viabilidad de este modelo en tiempo real, alcanzando un F1-score de 94.7% (Nasir et al., 2025).

Se ha estudiado, además de la detección de armas, la capacidad de esta arquitectura de detección para identificar ciertas acciones que despiertan sospechas y para contar personas, con el objetivo de comprobar su efectividad en situaciones de vigilancia por video. Se han creado sistemas que pueden detectar circunstancias como la presencia de armas o humo en lugares públicos, logrando niveles de precisión por encima del 95% y emitiendo alertas automáticas en tiempo real (Sudharson et al., 2023). De manera similar, se han introducido mejoras ligeras a la arquitectura base mediante módulos como soft-NMS y GSConv, obteniendo un 88.6% de precisión en videovigilancia de campus (Cheng et al., 2024). Además, se ha tratado el asunto de detectar objetos pequeños en imágenes aéreas captadas por drones. Para lograrlo, se ha sugerido el modelo SOD-YOLO, una versión perfeccionada de la estructura original que superó al YOLOv8l (Li et al., 2024) en un 7.7% del mAP@50. Esta perspectiva muestra que el detector sigue siendo eficaz, a pesar de los escenarios complicados que se observan en las imágenes aéreas obtenidas por drones.

El modelo YOLOv8, a diferencia de versiones previas como YOLOv5, YOLOv6 y YOLOv7, brinda progresos significativos en cuanto a la velocidad de inferencia, precisión y capacidad de generalización. Estas optimizaciones derivan de su estructura sin anclajes, la combinación más eficaz de atributos y los métodos de entrenamiento más sofisticados. Según otros análisis, YOLOv8 logró un 98% en F1-score y un 99% de precisión, lo que significa que superó ampliamente los resultados de YOLOv6 y YOLOv7 con 92% (Delwar et al., 2025). Por otro lado, el F1-score fue del 94.7% frente al 69% de YOLOv5 (Nasir et al., 2025). Estas pruebas respaldan el uso de esta arquitectura en esta investigación, ya que ofrece un equilibrio perfecto entre la eficacia y la velocidad de detección en tiempo real.

Es importante destacar que, aunque los avances más recientes muestran progresos notables en las métricas de precisión, la mayor parte de estos estudios no abordan el coste computacional y los requerimientos de hardware para la inferencia en tiempo real. El enfoque principal ha sido aumentar la precisión, frecuentemente suponiendo que se utilizan unidades de procesamiento gráfico de gama alta que no son económicamente factibles para la mayoría de las implementaciones de seguridad urbana o comercial. Esta disparidad entre la factibilidad práctica en hardware accesible y la exactitud teórica representa un reto crítico sin resolver. El principal problema que motivó este estudio radica

no solo en la limitada capacidad de los sistemas tradicionales de videovigilancia para detectar en tiempo real actividades sospechosas o la presencia de objetos peligrosos, sino también la ausencia de soluciones de inteligencia artificial validadas que mantengan un equilibrio entre una gran precisión y la eficiencia computacional requerida para su implementación en hardware económico. A partir de esta situación, surge la siguiente pregunta de investigación: ¿Cómo puede un sistema de videovigilancia inteligente basado en deep learning alcanzar al menos un 90% de precisión en la detección en tiempo real de personas, armas y actividades sospechosas?.

El objetivo general de este trabajo fue desarrollar un prototipo funcional de monitoreo inteligente capaz de detectar en tiempo real personas, armas y personas armadas utilizando un modelo basado en deep learning. De este se derivan los objetivos específicos: (1) construir un conjunto de datos personalizado que integre imágenes de personas, armas y personas armadas, aplicando técnicas de preprocesamiento y aumentación de datos; (2) entrenar y optimizar un modelo YOLOv8 mediante transfer learning para lograr altos niveles de precisión, recall, F1-score y mAP; y (3) integrar el modelo en un prototipo funcional de videovigilancia capaz de procesar video en tiempo real, emitir alertas basadas en reglas contextuales y validar su rendimiento mediante métricas cuantitativas y pruebas experimentales en entornos simulados y reales. Para alcanzar estos objetivos, se adoptó la metodología Cross-Industry Standard Process for Data Mining (CRISP-DM), estructurada en seis fases que guiaron de manera ordenada el desarrollo del sistema, desde la comprensión del problema y la preparación de datos hasta el modelado, la evaluación y la implementación final del prototipo.

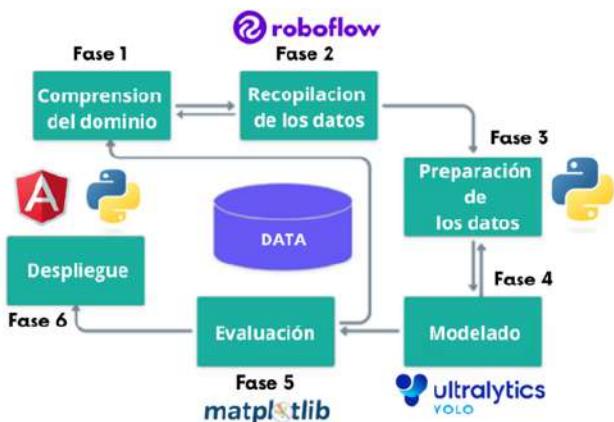
El artículo está organizado de la siguiente forma: en la Sección 2 se describen los materiales y métodos, en la Sección 3 se presentan los resultados y se discuten los hallazgos en relación con estudios anteriores, y en la Sección 4 se exponen las conclusiones y las futuras líneas de investigación.

## 2. Materiales y Métodos

El estudio fue llevado a cabo con la metodología CRISP-DM, que brinda una estructura ordenada para la creación de proyectos de análisis y modelaje de datos. Esta metodología fue escogida debido a su flexibilidad y a su habilidad para adecuarse a procesos iterativos, lo que posibilita hacer correcciones o volver atrás entre fases de acuerdo con los resultados alcanzados en cada etapa del desarrollo. El modelo CRISP-DM sigue siendo uno de los métodos más utilizados en iniciativas que se fundamentan en minería de datos y aprendizaje automático, gracias a su versatilidad y enfoque práctico, lo cual lo hace apropiado para contextos reales donde son esenciales la mejora permanente y la validación (Acuña-Cid et al., 2025).

En este trabajo, se utilizó CRISP-DM en seis etapas esenciales, como se observa en la Figura 1. Esta metodología sirvió de orientación para organizar el flujo del proyecto desde que se

entiende el dominio hasta que se valida el prototipo final. Durante estas etapas, se usaron herramientas específicas: Roboflow Universe para la recopilación y gestión del dataset; Python 3.11 para el procesamiento y preparación de datos; YOLOv8 en PyTorch para la fase de modelado; y Matplotlib para la visualización y evaluación de resultados. En última instancia, se emplearon Angular y Python (Flask) para desarrollar el frontend y el backend del prototipo, respectivamente; esto asegura una integración eficaz entre la interfaz de monitoreo y la inferencia del modelo.



**Figura 1.** Metodología CRISP-DM.

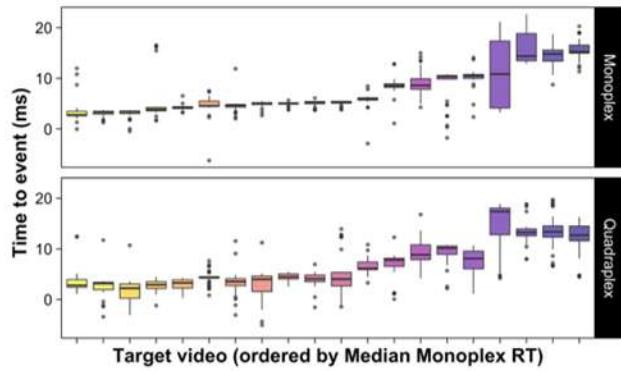
Fuente: Los autores.

El objetivo de esta etapa fue identificar las principales restricciones operativas de los sistemas tradicionales de videovigilancia y determinar la funcionalidad necesaria para crear un sistema inteligente de monitoreo. La fatiga, la distracción y las diferencias en la interpretación de los sucesos son ejemplos de factores humanos que impactan negativamente en el desempeño efectivo de la supervisión. De ahí surgió la necesidad de crear un sistema automatizado que tenga la capacidad de identificar armas y personas en tiempo real, aplicar normas contextuales para detectar conductas sospechosas y lanzar alertas automáticas. Esta perspectiva apunta a disminuir la dependencia de la supervisión manual, perfeccionar la exactitud en las detecciones y optimizar la capacidad de respuesta frente a incidentes en contextos urbanos y de seguridad privada.

Durante esta etapa se analizó el contexto operacional de la seguridad urbana y se documentaron fallas asociadas a la vigilancia sostenida: fatiga del operador, variabilidad en la interpretación de la escena y retrasos en la verificación de eventos, que conducen a falsos positivos y omisiones. Estudios recientes demuestran que, a medida que aumenta el número de pantallas a supervisar, el rendimiento del operador disminuye debido a la sobrecarga visual. Stainer et al. (2021) demostraron que, en entornos de videovigilancia con varias cámaras, los operadores presentan mayores tiempos de reacción y retrasos en la detección de eventos críticos. Como se observa en la Figura 2, el monitoreo



con una sola cámara genera tiempos de respuesta más cortos y precisos, mientras que el monitoreo de cuatro cámaras incrementa el retraso de detección de eventos sospechosos.



**Figura 2.** Variación del tiempo de respuesta ante tareas de videovigilancia

Fuente: Tomado de Stainer et al. (2021).

A partir de este diagnóstico se establecieron los siguientes criterios base para el desarrollo del sistema:

- R1: Identificación automática de armas y personas sin la intervención de seres humanos.
- R2: Disminución del tiempo de respuesta ante situaciones sospechosas.
- R3: Generación automática de alertas bajo condiciones preestablecidas.
- R4: Operación estable en tiempo real con un margen de error mínimo.

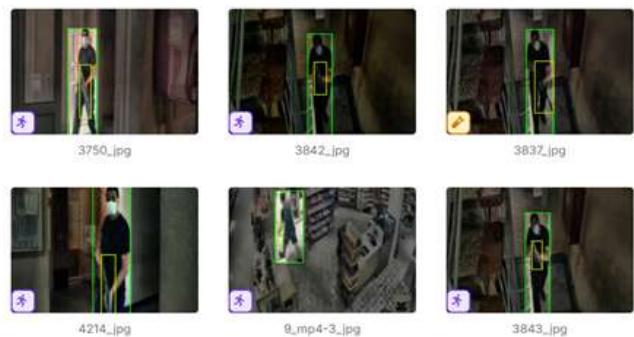
El resultado de esta etapa permitió definir las características del conjunto de datos, lo que sirvió de punto de partida para la siguiente fase.

#### Fase 2: Comprensión y recopilación de los datos

Los requisitos y criterios para el conjunto de datos que se requiere para entrenar el modelo de detección fueron determinados en esta etapa. La labor fundamental fue reunir imágenes que ilustraran la variedad de situaciones que un sistema de videovigilancia podría afrontar, teniendo en cuenta diferentes circunstancias, tales como las condiciones operativas, los ángulos de cámara y la iluminación.

Se eligieron imágenes que satisficieran tres requisitos principales: (1) la presencia evidente de las clases objetivo, es decir, persona, arma y persona armada; (2) una calidad

visual adecuada para anotar los objetos con precisión; y (3) la no existencia de distorsiones o marcas de agua que pudieran interferir con el entrenamiento. Se eliminaron las imágenes que tenían baja resolución, que se repetían o que presentaban etiquetas inconsistentes como se observa en la Figura 3.



**Figura 3.** Ejemplos de imágenes del dataset en Roboflow.

Fuente: Dataset completo de Erreyes (2025).

Para la construcción del dataset personalizado, se utilizaron tres repositorios de Roboflow Universe: el primero, “Armed Person Recognition Dataset” contiene más de 7,000 imágenes de personas armadas y fue elegido por su enfoque en detectar amenazas directas en entornos urbanos, el segundo tomado de “Personas con Armas Dataset” que incluye 6,600 imágenes de personas portando armas, ideal para identificar situaciones de riesgo y finalmente, “Weapon Detection Dataset” ofrece más de 12000 imágenes etiquetadas con diferentes tipos de armas, permitiendo al modelo reconocer objetos peligrosos.

Se llevó a cabo un análisis de las variables principales que influyen en el rendimiento de los modelos de detección en la videovigilancia con el fin de caracterizar objetivamente la diversidad del conjunto de datos. La Tabla 1 resume la distribución aproximada de imágenes en función de cada variable por separado, sobre un total de aproximadamente 8,500 imágenes recolectadas.

Esta distribución demuestra que el conjunto de datos incluye combinaciones complejas de condiciones operativas, lo cual permite que el modelo esté expuesto a situaciones que representan escenarios verdaderos de videovigilancia. La inclusión de situaciones difíciles como la baja iluminación, ángulos extremos, occlusiones severas asegura que el sistema entrenado sea capaz de generalizar frente a las condiciones cambiantes que suelen ocurrir en los ambientes comerciales y urbanos.

Para eliminar los duplicados y garantizar la coherencia en la

**Tabla 1.** Caracterización de la diversidad del dataset según variables operativas.  
Fuente: Los autores.

| Características     | Categorías         | Imágenes aproximadas |
|---------------------|--------------------|----------------------|
| Iluminación         | Alta               | 3,600                |
|                     | Media              | 3,200                |
|                     | Baja               | 1,700                |
| Ángulo de cámara    | Frontal (0-30°)    | 3,800                |
|                     | Oblicuo (30-60°)   | 3,000                |
|                     | Central (>60°)     | 1,700                |
| Contexto operativo  | Interior comercial | 4,500                |
|                     | Urbano exterior    | 2,500                |
|                     | Mixto/residencial  | 1,500                |
| Densidad de objetos | 1 persona/arma     | 3,400                |
|                     | 2-5 objetos        | 3,600                |
|                     | Más de 5 objetos   | 1,500                |
| Oclusiones          | Sin oclusión       | 4,000                |
|                     | Oclusión parcial   | 3,200                |
|                     | Oclusión severa    | 1,300                |

distribución de las clases, el conjunto de datos fue revisado manual y automáticamente antes del entrenamiento. Para prevenir sesgos de aprendizaje, se mantuvo un balance proporcional entre las categorías más importantes: persona, arma y persona armada. Asimismo, se examinó la congruencia entre las etiquetas y los objetos visibles para asegurar que el dataset final fuera de calidad.

El conjunto de datos recopilados fue organizado en tres subconjuntos principales: entrenamiento, validación y prueba,

**Tabla 2.** Composición del Conjunto de Datos.  
Fuente: Los autores.

| Clase          | Entrenamiento | Validación | Prueba |
|----------------|---------------|------------|--------|
| Persona        | 8,176         | 857        | 893    |
| Arma           | 9,018         | 888        | 964    |
| Persona armada | 7,030         | 758        | 758    |

con el fin de garantizar una evaluación adecuada. La Tabla 2 muestra la distribución de las instancias por clase, evidenciando un balance entre categorías de persona, arma y persona armada.

La división se realizó utilizando la proporción 8:1:1 (80% entrenamiento, 10% validación y 10% prueba), asegurando la distribución equilibrada de clases en cada subconjunto. Esta estructura fue seleccionada en base a referencias metodológicas previas, como la proporción 7:2:1 de Ni et al. (2024) aplicados en entornos subterráneos y el esquema 3:1 (75% entrenamiento y 25% validación) del trabajo de Delwar et al. (2025) para proyectos en entornos de agricultura. De esta manera, optar por la proporción 8:1:1 nos dispone de suficientes datos para

el entrenamiento, con el fin de reducir el riesgo del underfitting (subajuste), y al mismo tiempo reservar datos de validación y prueba que actúan frente al overfitting (sobreajuste) garantizando una evaluación imparcial del rendimiento.

Este análisis se basa en los principios de la IA responsable, que valoran ante todo la transparencia, la equidad y la protección de datos. Al elegir y preparar el conjunto de datos, se pusieron en práctica acciones destinadas a salvaguardar la privacidad y reducir los peligros de sesgos algorítmicos. Según Castro-Paredes et al. (2025), es preciso incluir la protección de datos desde el diseño original del sistema y restringir la recopilación de datos a lo que sea estrictamente necesario. En este contexto, se eliminaron los metadatos identificativos y se comprobó que las imágenes fueran obtenidas de repositorios públicos como Roboflow Universe, los cuales tienen licencias abiertas para la investigación y el uso académico. De acuerdo con los términos establecidos por sus creadores, se utilizaron las colecciones de datos elegidas sin alterar el contenido original más allá de las operaciones de aumento. Se respetaron los principios de privacidad y el uso ético de datos al no incluir información personal ni metadatos que pudieran revelar la identidad de personas.

Además, se abordó la advertencia de Alvarado & Villavicencio (2024) sobre cómo los algoritmos de aprendizaje automático pueden heredar y aumentar los sesgos que ya existen en los datos con los que son entrenados. Para evitar la discriminación algorítmica, se documentó la diversidad del conjunto de datos en cuanto a iluminación, ángulos de cámara y contextos operativos, como se muestra en la Tabla 1. Esta caracterización permitió confirmar que el conjunto de datos no tuviera una sobrerepresentación de condiciones particulares que pudieran inducir sesgos en el modelo final. El conjunto de datos se empleó solamente para validar y desarrollar modelos de detección orientados a la seguridad pública y privada.

### Fase 3: Preparación de los datos

Después de etiquetar el conjunto de datos, se empezó a preparar el dataset para el entrenamiento con la finalidad de optimizar la habilidad de generalización ante cambios comunes en videovigilancia como las variaciones en la iluminación, en los ángulos de cámara, oclusiones parciales y escalas disímiles. La elección de las técnicas de aumentación se llevó a cabo teniendo en cuenta la problemática concreta que son las personas, armas y personas armadas, dándole prioridad a los cambios que aporten alteraciones realistas sin menoscabar el significado de las escenas ni la forma exterior de objetos pequeños.

Las modificaciones geométricas como la perspectiva, rotación, traslación, escala, shear y mosaico persiguen una solidez tanto espacial como multi-escala; los cambios en HSV se ocupan de la variabilidad luminosa propia de interiores y exteriores; el mixup y el copy-paste atenuan los límites de decisión y elevan las co-ocurrencias, lo que disminuye el sobreajuste. Las probabilidades se establecieron para mantener un equilibrio entre la fidelidad y la diversidad: son elevadas cuando la transformación



**Tabla 3.** Técnicas de aumento de datos y sus parámetros aplicados.

Fuente: Los autores.

| Técnica           | Objetivo                                      | Probabilidad          | Efecto esperado  |
|-------------------|---|-----------------------|--|
| Mosaic            | Diversidad de contextos y escalas simultáneas | $p = 0.5$             | Mejor detección multi-escala; menor sobreajuste.               |
| MixUp             | Suavizar fronteras de decisión                | $p = 0.1$             | Mayor robustez a occlusiones/ruido; Clases más separables.     |
| Volteo horizontal | Invariancia a orientación lateral             | $p = 0.5$             | Generalización a distintas direcciones de movimiento.          |
| Rotación          | Tolerancia a inclinaciones leves              | Hasta 10° (aleatoria) | Estabilidad ante cámaras no niveladas.                         |
| Traslación        | Robustez a encuadres variables                | Hasta 20%             | Menor sensibilidad a reencuadres del objeto.                   |
| Escalado          | Robustez multi-escala                         | Hasta 50%             | Mejora detección de objetos pequeños/grandes.                  |
| Shear             | Variación geométrica sutil                    | 2°                    | Mejora frente a distorsiones por perspectiva.                  |
| Perspectiva leve  | Simular cámaras en altura                     | Aleatoria, leve       | Mejor desempeño en CCTV montado en techos/paredes.             |
| Ajuste HSV        | Variabilidad de iluminación/color             | Deltas aleatorios     | Robustez a interiores/exteriores y sombras.                    |
| Copy-Paste        | Aumentar co-ocurrencias y densidad            | $p = 0.2$             | Más ejemplos “persona armada”; mejora recall sin sobrerotular. |

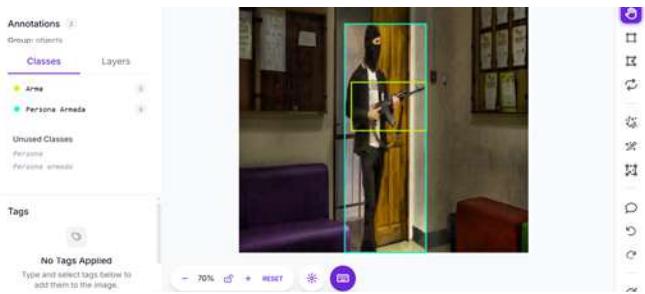
reproduce condiciones comunes como el volcado horizontal, y moderadas o bajas cuando hay peligro de agregar ruido principalmente en mixup o copy-paste. La Tabla 3 resume cada técnica de aumento de datos aplicada, su propósito, los parámetros probabilísticos y el efecto esperado sobre la generalización del modelo.

Estas estrategias se aplicaron para equilibrar las clases y aumentar la diversidad del conjunto de datos, siguiendo enfoques similares a los reportados por Espinoza et al. (2025), quienes evidencian que el incremento controlado de la variabilidad de las muestras contribuye a una mejor generalización de los modelos.

Para garantizar la coherencia y compatibilidad con el modelo elegido, se realizaron las siguientes acciones preliminares:

- Todas las imágenes fueron ajustadas a una resolución estándar de  $640 \times 640$  píxeles, para mantener un equilibrio entre precisión y velocidad de inferencia.

- Se verificó que las imágenes estuvieran en formato jpg, y las etiquetas asociadas en archivos .txt con coordenadas normalizadas de los bounding boxes.



**Figura 4.** Interfaz de anotación de clases en Roboflow.

Fuente: Los autores.

El proceso de anotación de datos se realizó a través de la plataforma Roboflow Universe, la cual facilitó la visualización y edición gráfica de las etiquetas mediante una interfaz que permite delimitar los objetos con bounding boxes de distintos colores, como se observa en la Figura 4.

Las imágenes provinieron de los repositorios seleccionados fueron previamente etiquetados por sus autores originales; sin embargo, fue necesario revisar y ajustar manualmente algunas anotaciones, corrigiendo nombres de clases y coordenadas de los bounding boxes para garantizar la coherencia y precisión de las etiquetas. Este proceso permitió mantener la consistencia entre las clases “persona”, “arma” y “persona armada”, asegurando un formato estandarizado compatible con YOLOv8.

Se llevó a cabo una revisión mixta, automática y manual de la totalidad de los datos con el fin de asegurar que las etiquetas fueran coherentes y que las clases estuvieran correctamente asignadas. Para detectar errores de formato, clases no definidas y coordenadas que exceden el rango normalizado, se utilizaron scripts de validación en la verificación automática. Además, se realizó una revisión manual a través de la plataforma Roboflow, que permite ver y editar directamente las etiquetas en cada imagen. Esta herramienta posibilitó verificar la exactitud de los bounding boxes y rectificar potenciales inconsistencias o superposiciones en las clases objetivo. En total, se examinó manualmente alrededor del 10 % de los datos totales para garantizar la coherencia semántica y visual de las anotaciones antes de que el modelo comenzara su entrenamiento.

#### Fase 4: Modelado

En esta fase se realizó el entrenamiento del modelo YOLOv8 usando el conjunto de datos preparado en su totalidad. Se configuraron los hiperparámetros más importantes como el número de épocas, el tamaño de lote y la tasa de aprendizaje, el fin de esto es la optimización del proceso de aprendizaje y no tener un sobreajuste innecesario. Se implementaron métodos de validación y modificaciones dinámicas en la tasa de aprendizaje a lo largo del entrenamiento para mejorar el proceso de detección y estabilizar el aprendizaje del modelo.

Se utilizó el modelo YOLOv8 en dos versiones para el entrenamiento la versión nano (YOLOv8n) y la versión small (YOLOv8s). La versión estándar YOLOv8 no se seleccionó porque YOLOv8s la supera en velocidad y rendimiento. Las dos versiones fueron elegidas con el propósito de valorar la eficacia en cuanto a rapidez de inferencia y precisión, buscando así hallar la mejor relación entre eficiencia y exactitud para un despliegue viable en hardware accesible y de bajo costo.

- YOLOv8n:** Se optó por esta versión más liviana al principio porque necesita menos recursos computacionales, lo que la hace perfecta para aplicaciones en tiempo real que demandan un uso reducido de memoria y capacidad de procesamiento.

- YOLOv8s:** Es un modelo más sólido que logra optimizar la capacidad de generalización y precisión, a la vez que mantiene una latencia apropiada para labores de videovigilancia en sistemas con recursos computacionales limitados.

El modelo se desarrolló, entrenó y evaluó en un entorno local de alto rendimiento, equipado con una GPU NVIDIA RTX 2050 4 GB VRAM, adecuada para tareas de detección en tiempo real mediante deep learning. El sistema operativo utilizado fue Windows 11 Pro, con 16 GB de memoria RAM y un almacenamiento SSD de 1 TB, lo que permitió una ejecución fluida del entrenamiento y validación del modelo YOLOv8s. El lenguaje principal de programación fue Python 3.11, utilizando la librería Ultralytics YOLOv8 basada en PyTorch. Todo el desarrollo fue gestionado en el entorno de trabajo Visual Studio Code, que facilitó la integración de scripts, pruebas y gestión de

dependencias durante el proceso de experimentación.

La configuración de hiperparámetros constituye un elemento clave para el rendimiento del modelo, ya que determina tanto

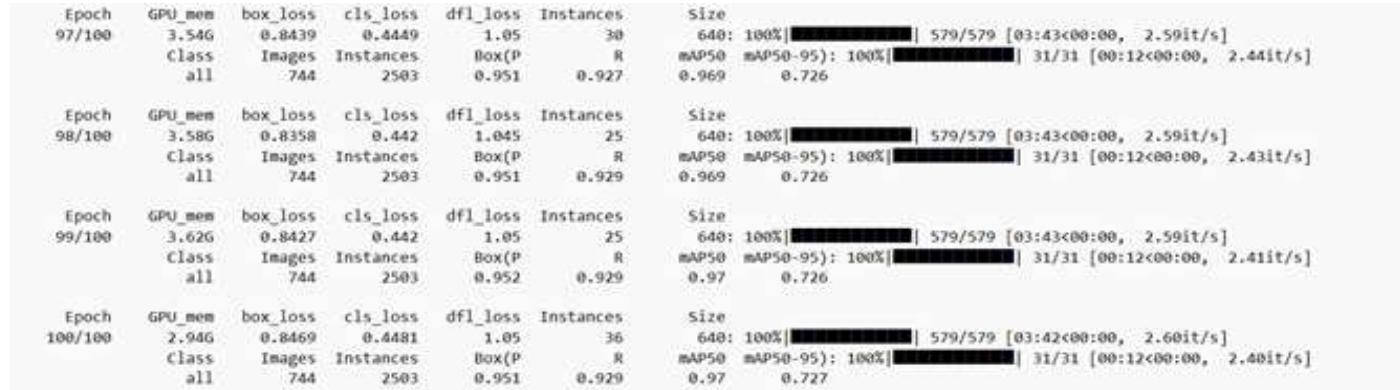
**Tabla 4.** Configuración de Hiperparámetros.

Fuente: Los autores.

| Parámetro             | Prueba 1<br>(YOLOv8n) | Prueba 2<br>(YOLOv8s) | Prueba 3<br>(YOLOv8s) | Prueba 4<br>(YOLOv8s) |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Modelo base           | YOLOv8n               | YOLOv8s               | YOLOv8s               | YOLOv8s               |
| Épocas                | 50                    | 75                    | 100                   | 100                   |
| Batch size            | 16                    | 8                     | 12                    | 12                    |
| Learning rate (lr0)   | 0.001                 | 0.001                 | 0.0008                | 0.0005                |
| Learning rate (lrf)   | 0.0002                | 0.0002                | 0.00016               | 0.0001                |
| Warmup epochs         | 3                     | 3                     | 5                     | 5                     |
| Weight decay          | 0.0005                | 0.0005                | 0.0005                | 0.0005                |
| Mosaic                | 0.3                   | 0.4                   | 0.5                   | 0.5                   |
| Mixup                 | 0.0                   | 0.05                  | 0.1                   | 0.1                   |
| Copy-paste            | 0.0                   | 0.1                   | 0.15                  | 0.2                   |
| Cosine LR             | No                    | Sí                    | Sí                    | Sí                    |
| Optimizer             | SGD                   | AdamW                 | AdamW                 | AdamW                 |
| Caché                 | RAM                   | RAM                   | Disk                  | Disk                  |
| AMP (Mixed Precision) | Sí                    | Sí                    | Sí                    | Sí                    |

la estabilidad como la capacidad de generalización del sistema. Los valores utilizados en esta investigación están especificados en la Tabla 4; fueron seleccionados para asegurar un rendimiento equilibrado del modelo y una inferencia apropiada en tiempo real.

El entrenamiento del modelo se realizó en el entorno local descrito anteriormente, aprovechando la GPU para acelerar los cálculos de propagación hacia adelante y hacia atrás. En la Figura 5 se muestra un fragmento del registro de un entrenamiento, donde se



**Figura 5.** Registro del proceso de entrenamiento del modelo YOLOv8s durante 100 épocas.

Fuente: Los autores.



evidencian las perdidas progresivas. El flujo de trabajo incluyó:

- Inicialización del modelo con pesos preentrenados (transfer learning) aprovechando el conocimiento previo en tareas de detección general y ajustarlo al conjunto de datos específico (fine-tuning).
- Carga del dataset preprocesado y organizado en las carpetas train/, val/ y test/, utilizando el módulo YOLODataset de la librería Ultralytics.
- Ejecución de ciclos de entrenamiento durante 100 épocas, aplicando el algoritmo AdamW para actualizar los parámetros de la red.
- Monitoreo en tiempo real de métricas clave como:
  - Box loss: error asociado a la precisión de las coordenadas de los bounding boxes.
  - Class loss: error en la clasificación de las detecciones.
  - Objectness loss: error en la predicción de la existencia de objetos.

El entrenamiento se realizó con un tiempo total aproximado de entre 4 a 6 horas. El tiempo medio por época fue de cerca de 4 minutos. La resolución de las imágenes empleadas fue  $640 \times 640$  píxeles, y el consumo de memoria GPU osciló entre 2.8 GB y 3.63 GB, dependiendo del tamaño de los lotes y las configuraciones particulares de cada versión.

#### Fase 5: Evaluación

La evaluación de los modelos se realizó utilizando un subconjunto de pruebas que equivale al 10% del conjunto de datos original, conformado por imágenes no utilizadas durante la fase de entrenamiento ni validación. Este conjunto de pruebas incluyó:

- Aproximadamente 700 imágenes distribuidas proporcionalmente entre las clases persona, arma y persona armada.
- Escenarios variados con diferente iluminación, ángulos de cámara y densidad de objetos, representando condiciones operativas reales.

La Tabla 5 presenta un resumen de estas métricas, que incluyen la precisión y el recall como indicadores básicos, el F1-score como equilibrio entre ambas y los valores de mAP@50 y mAP@50-95 como referencias utilizadas para cuantificar la calidad de detección bajo diferentes umbrales.

Las herramientas de YOLOv8 calculan de manera automática

**Tabla 5.** Métricas de Evaluación.

Fuente: Los autores.

| Métrica               | Descripción  |
|-----------------------|--|
| Precisión (Precision) | Proporción de verdaderos positivos sobre todas las predicciones positivas realizadas.                                |
| Recall (Sensibilidad) | Proporción de verdaderos positivos correctamente identificados sobre todas las instancias reales.                    |
| F1-Score              | Media armónica entre precisión y recall, útil para medir el balance entre ambas métricas.                            |
| mAP@50                | Media de precisión considerando un umbral de IoU=0.5, estándar en la comunidad de visión por computador.             |
| mAP@50-95             | Media de precisión sobre múltiples umbrales de IoU (0.5 a 0.95 con pasos de 0.05), para una evaluación más rigurosa. |

las métricas de evaluación mientras se validaba el modelo. Los resultados de estas métricas se obtienen del entrenamiento y se guardan en la carpeta de resultados del conjunto de datos de test. Las funciones de metrics.py, que son parte del framework de YOLOv8, se emplean para calcular las métricas. Las mismas que se adquieren directamente de las estimaciones hechas en el conjunto de datos de prueba, sin el uso de herramientas externas.

#### Fase 6: Despliegue

La fase de despliegue implicó la incorporación del modelo YOLOv8s en un prototipo funcional que era capaz de funcionar en tiempo real. Este sistema tiene la habilidad de procesar flujos de video, detectar objetos y lanzar alertas en tiempo real al descubrir acciones sospechosas como la presencia de individuos armados en áreas restringidas.

El modelo que se entrenó fue extraído del ambiente de entrenamiento y añadido a un prototipo de sistema de vigilancia por video, que fue creado con OpenCV y Python, empleando las capas siguientes:

**Capa de datos:** Cámaras IP o locales que, por medio del protocolo RTSP/IP, graban video en tiempo real y transmiten fotogramas de video para su procesamiento.

**Capa de Backend:**

- **Detección de objetos:** Para la detección en tiempo real, YOLOv8 es el componente esencial.
- **Servicios RESTful:** Flask v2.0 es la herramienta que se utiliza para administrar los servicios de backend; esta posibilita que el modelo de detección y el frontend se comuniquen entre sí.

- Comunicación en tiempo real:** Para garantizar una comunicación eficaz entre el frontend y el backend, se emplean WebSockets.

- Envío de alertas:** Para la transmisión de alertas de amenaza o intrusión a los usuarios, se emplea la API v3 de Mailgun.

#### Capa de Frontend:

- Interfaz de usuario interactiva: Creada con Angular v12, posibilita la visualización en tiempo real del estado del sistema y de las detecciones.

- Recepción en tiempo real: Las alertas y la información de las cámaras llegan al frontend a través de WebSockets.

- Estructura y diseño UI: Se utiliza HTML/CSS para el diseño de la interfaz, con el fin de brindar una experiencia eficaz y amigable.

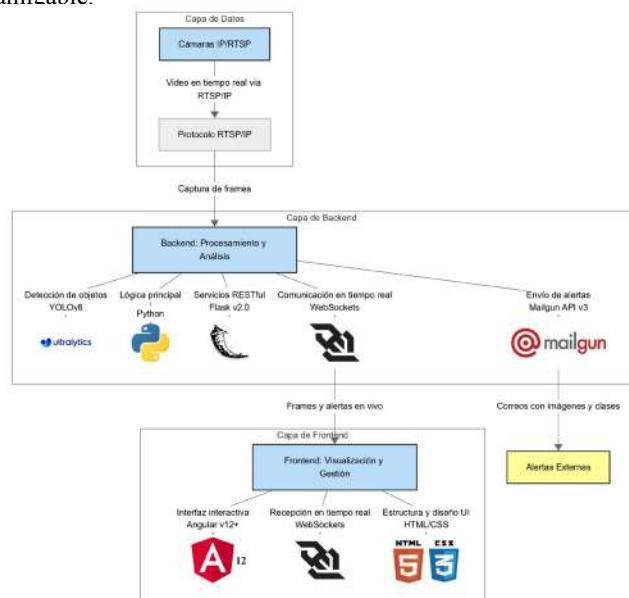


Figura 6. Arquitectura tecnológica del Prototipo.

Fuente: Los autores.

#### Diagrama de arquitectura tecnológica

El diagrama de la arquitectura tecnológica del sistema, que refleja las capas del software, la comunicación entre los componentes y el flujo de datos se ilustra en la Figura 6.

La Figura 7 presenta en detalle la arquitectura del prototipo, que muestra el funcionamiento del flujo de integración del modelo, desde que se captura el video en tiempo real hasta que se procesa con YOLOv8s y se clasifican las amenazas. Los resultados se almacenan en la parte trasera del sistema, se muestran en una página web y, si se identifican riesgos, el sistema notifica automáticamente al usuario.

#### Requerimientos de hardware y tiempo promedio de respuesta

El sistema de videovigilancia en tiempo real está creado para

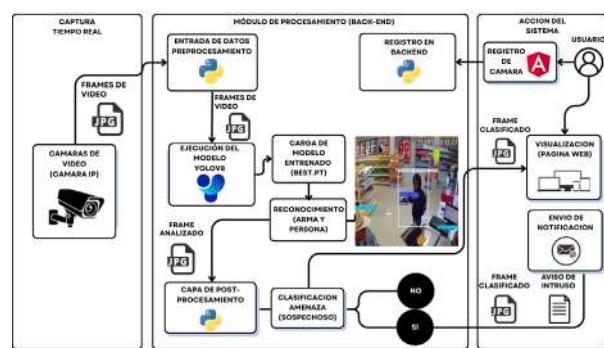


Figura 7. Arquitectura del Prototipo.

Fuente: Los autores.

funcionar con eficiencia elevada. Los requisitos mínimos de hardware y el promedio del tiempo de respuesta se describen a continuación:

#### Hardware:

- GPU:** GeForce GTS 450/Radeon HD 6850, apropiada para el procesamiento de los modelos de detección en tiempo real.
- Memoria RAM:** Se recomienda 8 GB de RAM para optimizar la carga de trabajo del sistema.
- CPU:** Un procesador Intel Core i3 o similar para asegurar el procesamiento eficiente de los flujos de video.

#### Tiempo promedio de respuesta global:

- Captura y procesamiento de frames:** El tiempo total por imagen es de aproximadamente 30 ms, lo cual incluye el preprocesamiento de 0.5 ms, la inferencia de 22 ms y el posprocesamiento de 7.5 ms.

#### Pruebas de usabilidad y validación funcional

Se llevaron a cabo pruebas de usabilidad y validación funcional, que integraron análisis cualitativos y mediciones cuantitativas, con el fin de determinar cuán eficiente es el sistema para detectar amenazas, sobre todo en cuanto a la identificación de personas armadas. Para evaluar la eficacia de las alertas en tiempo real, la facilidad de uso y la habilidad del sistema para mostrar imágenes de varias cámaras al mismo tiempo, los usuarios utilizaron la interfaz. Además, se llevaron a cabo exámenes funcionales con diez videos representativos de distintos contextos operativos que contenían variaciones en iluminación, ángulos de cámara y presencia de personas armadas y desarmadas. El objetivo era evaluar el desempeño del sistema en relación con falsos positivos y falsos negativos. Se utilizaron medidas estándar de visión por computadora, tales como precisión, mAP@50, mAP@50–95, recall y F1-score para medir su rendimiento. Los resultados obtenidos posibilitaron la confirmación de que el prototipo es robusto en situaciones adversas y controladas.

#### Consideraciones éticas y legales del sistema

El uso de tecnologías de videovigilancia inteligente conlleva responsabilidades legales y éticas que necesitan ser atendidas



explícitamente. Se utilizaron videos que se obtuvieron de fuentes públicas disponibles en plataformas como YouTube y sitios web de medios noticiosos, los cuales correspondían a circunstancias reales de robos e incidentes de seguridad registrados por cámaras de videovigilancia, durante la fase de validación experimental. Estos materiales fueron empleados solamente con propósitos académicos y de investigación, sin procesar datos biométricos identificables ni guardar información personal. Según Castro-Paredes et al. (2025), en Ecuador, la Ley Orgánica de Protección de Datos Personales define principios esenciales para el manejo de datos personales, como son la transparencia, el consentimiento informado y la reducción de datos.

### 3. Resultados y Discusión

#### Resultados Cuantitativos

Tras completar el proceso de entrenamiento y validación de las distintas configuraciones de YOLO, fue necesario establecer un criterio objetivo para determinar cuál de los modelos presentaba el mejor desempeño. Para ello, se recurrió a un conjunto de métricas ampliamente utilizadas en visión por computadora, que permiten medir no solo la precisión de las detecciones, sino también la capacidad del sistema para reconocer correctamente las clases definidas bajo diferentes condiciones. Estas métricas proporcionan una base cuantitativa sólida para comparar los resultados obtenidos en cada prueba y respaldar la selección final del modelo que será implementado en el prototipo de monitoreo inteligente.

Se evaluó el rendimiento de los modelos usando métricas estándar en visión por computador, las cuales fueron calculadas con base en las matrices de confusión producidas en cada contexto de prueba:

**Precisión (Precision):** calcula el porcentaje de verdaderos positivos en relación con todas las predicciones positivas.

$$P = TP / (TP + FP) \quad (1)$$

**Sensibilidad (Recall):** señala la relación entre la cantidad de verdaderos positivos que se han identificado correctamente y el total de casos reales.

$$R = TP / (TP + FN) \quad (2)$$

**F1-Score:** es la media armónica de las medidas de precisión y de recall.

$$F1 = 2 \times (P \times R) / (P + R) \quad (3)$$

**mAP@50:** promedio de precisión que tiene en cuenta un límite de IoU igual a 0.5.

$$mAP@50 = (1/N) \times \sum AP_i \text{ con } IoU \geq 0.5 \quad (4)$$

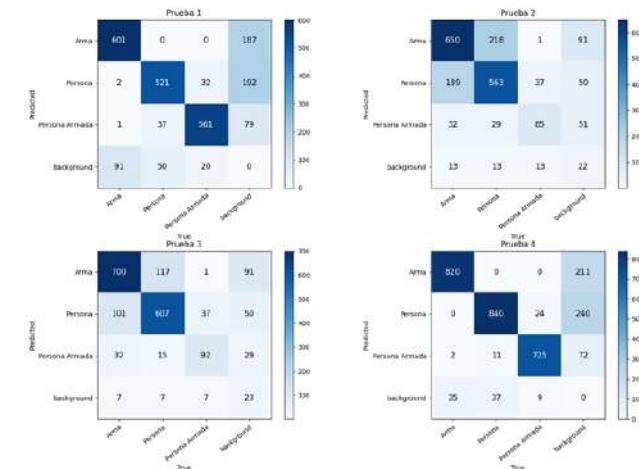
**mAP@50-95:** media de la precisión en varios límites de IoU.

$$mAP@50-95 = (1/10) \times \sum AP_i \text{ con } 0.5 \leq IoU \leq 0.95, \text{ paso } 0.05 \quad (5)$$

#### Donde:

- TP (True Positives): cantidad de casos que se ha clasificado correctamente como positivos.
- FP (False Positives): cantidad de casos que han sido mal clasificados como positivos.
- FN (False Negatives): cantidad de casos positivos que el modelo no fue capaz de detectar.
- TN (True Negatives): cantidad de casos que se clasificaron de manera correcta como negativos.

Se evaluaron las 4 pruebas de YOLO con el uso de la matriz



**Figura 8.** Matrices de evaluación de las Pruebas de YOLO.  
(Prueba 1) Primera imagen; (Prueba 2) Segunda imagen;  
(Prueba 3) Tercera imagen; (Prueba 4) Cuarta imagen.

Fuente: Los autores.

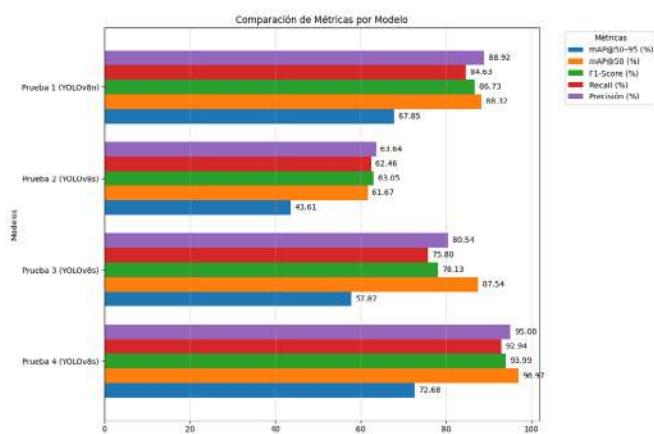
de confusión de cada uno respectivamente como se observa en la Figura 8. Estas matrices permitieron observar la evolución del desempeño de los modelos desde la primera prueba con YOLOv8n hasta la configuración final que fue con YOLOv8s. La Tabla 6 y la Figura 9 presentan el rendimiento comparativo de los modelos en las métricas F1-score, recall, precisión y mAP. La Prueba 4 YOLOv8s alcanzó los valores más altos, con una precisión del 95.08 % y un mAP@50 de 96.97 %, demostrando así su habilidad superior para identificar objetos con gran exactitud. Por otro lado, los resultados de las pruebas 2 y 3 mostraron rendimientos moderados, lo que evidencia el impacto del tamaño del conjunto de datos y la configuración de hiperparámetros. La

Prueba 1 YOLOv8n demostró una estabilidad inferior, lo que confirma que YOLOv8s brinda un mayor balance entre precisión

**Tabla 6.** Comparativa de rendimiento de los modelos entrenados.

Fuente: Los autores.

| Modelo             | Precisión (%) | Recall (%) | F1-Score (%) | mAP@50 (%) | mAP@50–95 (%) |
|--------------------|---------------|------------|--------------|------------|---------------|
| Prueba 1 (YOLOv8n) | 88.92         | 84.63      | 86.73        | 88.32      | 67.85         |
| Prueba 2 (YOLOv8s) | 63.64         | 62.46      | 63.05        | 61.67      | 43.61         |
| Prueba 3 (YOLOv8s) | 80.54         | 75.80      | 78.13        | 87.54      | 57.87         |
| Prueba 4 (YOLOv8s) | 95.08         | 92.94      | 93.99        | 96.97      | 72.68         |



**Figura 9.** Comparación de métricas de rendimiento entre modelos YOLOv8n y YOLOv8s.

Fuente: Los autores.

y velocidad. Los resultados, en general, corroboran que la prueba final YOLOv8s es la más apropiada para el seguimiento inteligente, ofreciendo el mejor equilibrado entre alta precisión y la eficiencia requerida para hardware de bajo costo.

La Figura 10 muestra la interfaz del sistema de videovigilancia inteligente en funcionamiento, se puede observar el flujo de video en tiempo real donde resaltan las etiquetas de “persona”, “arma” y “persona armada” siendo estos los objetos que están detectados por el modelo, también se encuentra el panel de las “detecciones activas”, la cual resume cada hallazgo con su porcentaje precisión y sus coordenadas exactas. Para evaluar el prototipo de monitoreo inteligente, se realizaron diez pruebas experimentales utilizando videos grabados en interiores y exteriores de locales comerciales, en diferentes condiciones de iluminación y con variación en los ángulos de cámara. En

cada escenario se representaron situaciones con personas, armas y personas armadas, con el fin de verificar la capacidad de detección y reconocimiento en un entorno real. El modelo seleccionado, YOLOv8s en su configuración final (Prueba 4), procesó los videos en tiempo real y los resultados obtenidos fueron contrastados con las predicciones esperadas a partir de la fase de entrenamiento, lo que permitió medir con precisión su desempeño en escenarios de seguridad no controlados.

Para profundizar en el rendimiento del sistema, se elaboró un análisis cuantitativo de errores a partir de los diez videos de prueba. En la Tabla 7 se puede verificar cada caso en donde se registraron los valores de Falsos Positivos (FP), Falsos Negativos (FN), Verdaderos Positivos (TP) y Verdaderos Negativos (TN) para cada clase de Persona, Arma y Persona armada, tienen diez escenarios de prueba.

- **Clase Persona:** alcanzó valores altos de TP en todos los

**Tabla 7.** Análisis cuantitativo de los videos por clase.

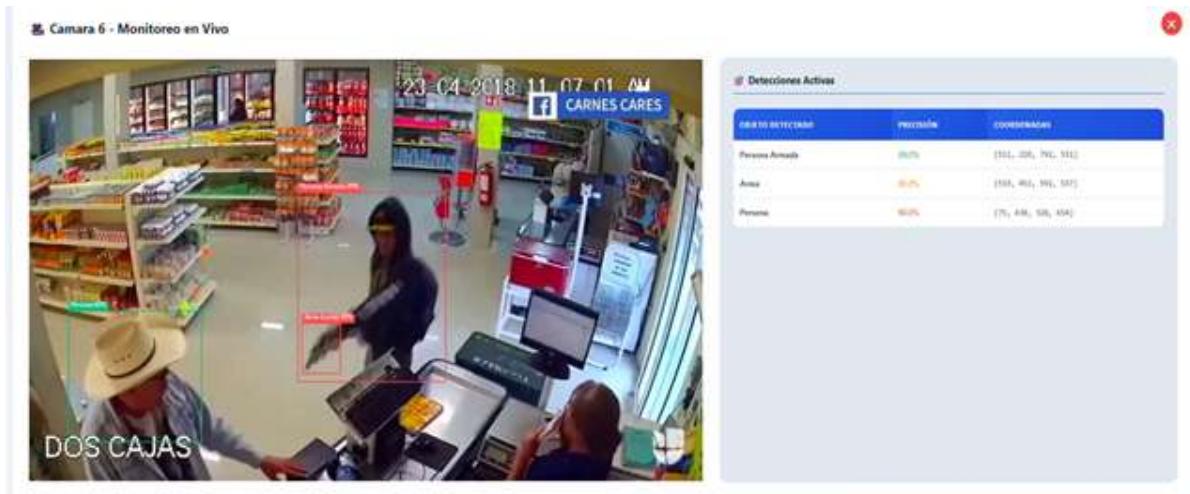
Fuente: Los autores.

| Video | Clase          | FP | FN | TP | TN |
|-------|----------------|----|----|----|----|
| 1     | Persona        | 2  | 1  | 27 | 70 |
|       | Arma           | 1  | 2  | 8  | 89 |
|       | Persona armada | 0  | 1  | 9  | 90 |
| 2     | Persona        | 3  | 0  | 28 | 69 |
|       | Arma           | 2  | 1  | 10 | 87 |
|       | Persona armada | 1  | 2  | 8  | 89 |
| 3     | Persona        | 1  | 2  | 26 | 71 |
|       | Arma           | 0  | 3  | 9  | 88 |
|       | Persona armada | 2  | 1  | 9  | 88 |
| 4     | Persona        | 2  | 1  | 27 | 70 |
|       | Arma           | 1  | 1  | 11 | 87 |
|       | Persona armada | 0  | 2  | 8  | 90 |
| 5     | Persona        | 1  | 1  | 28 | 70 |
|       | Arma           | 0  | 2  | 10 | 89 |
|       | Persona armada | 1  | 1  | 9  | 89 |
| 6     | Persona        | 2  | 0  | 29 | 69 |
|       | Arma           | 1  | 2  | 9  | 88 |
|       | Persona armada | 0  | 1  | 10 | 89 |
| 7     | Persona        | 3  | 2  | 25 | 71 |
|       | Arma           | 2  | 3  | 7  | 88 |
|       | Persona armada | 1  | 2  | 8  | 89 |
| 8     | Persona        | 1  | 1  | 30 | 68 |
|       | Arma           | 1  | 2  | 9  | 88 |
|       | Persona armada | 0  | 1  | 10 | 89 |
| 9     | Persona        | 2  | 3  | 26 | 70 |
|       | Arma           | 3  | 2  | 8  | 87 |
|       | Persona armada | 2  | 1  | 9  | 88 |
| 10    | Persona        | 1  | 1  | 28 | 70 |
|       | Arma           | 1  | 1  | 10 | 88 |
|       | Persona armada | 0  | 1  | 9  | 90 |

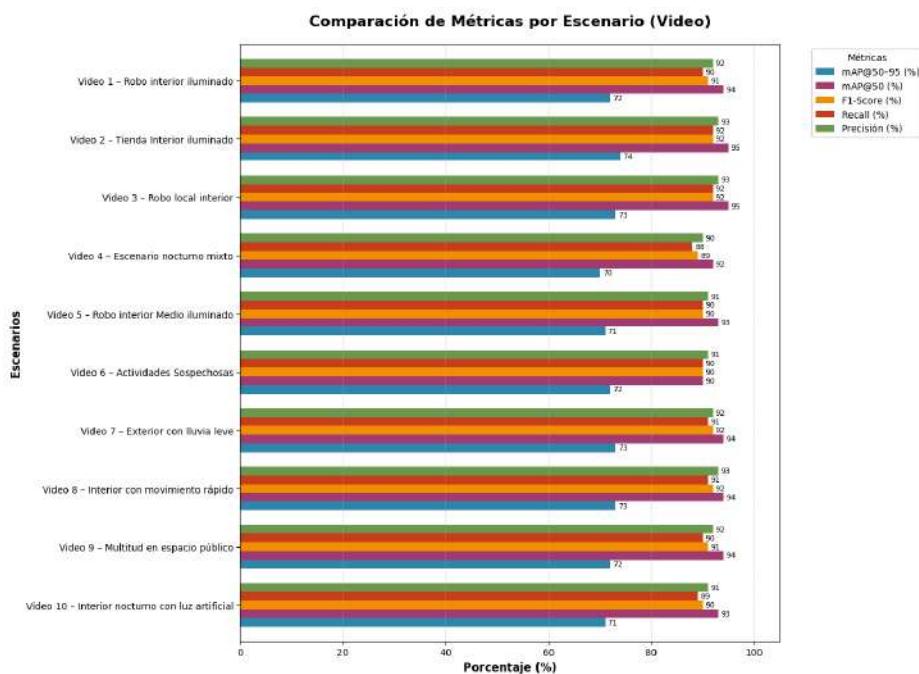
escenarios (26–30), con FN bajos (1–3), confirmando que el modelo mantiene su robustez para identificar individuos.

- **Clase Arma:** fue la más sensible a variaciones ambientales, con algunos FP (1–2) y FN (2–4), lo que evidencia confusiones con objetos similares en baja luz o ángulos difíciles.





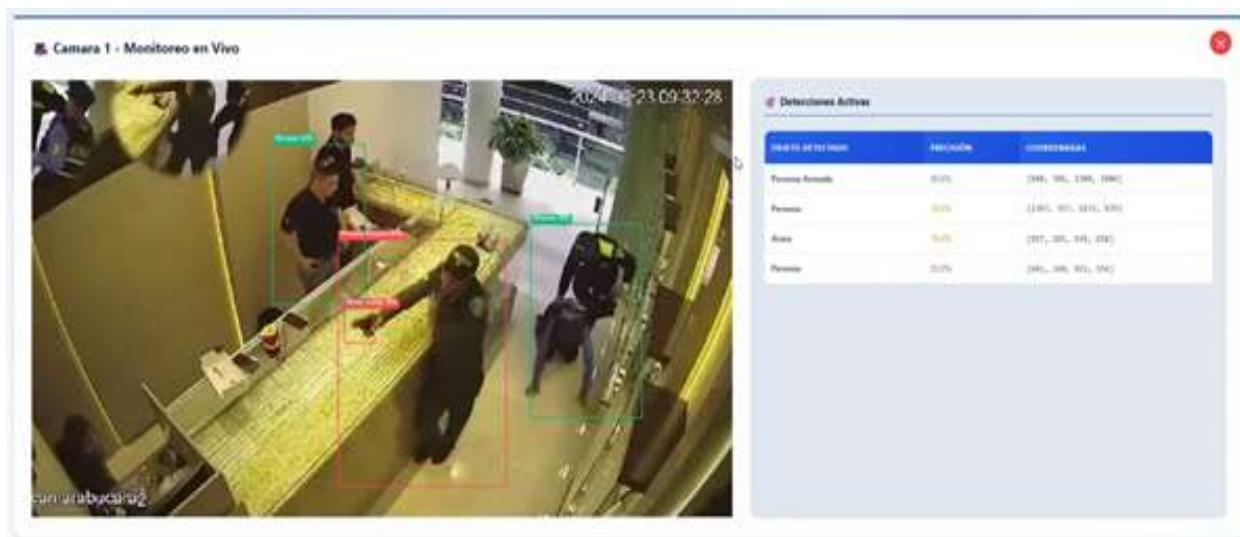
**Figura 10.** Interfaz de prototipo del sistema de vigilancia inteligente.  
Fuente: Los autores.



**Figura 11.** Comparación de métricas por escenario de validación del prototipo.  
Fuente: Los autores.

**Tabla 9.** Matriz de confusión detallada del modelo final  
Fuente: Los autores.

| Matriz de confusión  | Predicción Persona | Predicción Arma | Predicción Persona armada |
|----------------------|--------------------|-----------------|---------------------------|
| Real: Persona        | 840                | 18              | 32                        |
| Real: Arma           | 40                 | 211             | 15                        |
| Real: Persona armada | 25                 | 22              | 210                       |



**Figura 12.** Desempeño del modelo en entorno controlado (Video 1 – Interior iluminado).

Fuente: Los autores.



**Figura 13.** Desempeño del modelo en condiciones adversas (Video 4 – Escenario nocturno mixto).

Fuente: Los autores.

**Clase Persona armada:** se mantuvo en un rango de 8–10 TP por escenario, aunque con FN asociados a armas parcialmente ocultas y FP mínimos.

Los resultados obtenidos en cada escenario se muestran en una matriz donde se evalúan las métricas establecidas previamente, donde se realiza el cálculo de la diferencia entre lo previsto y lo obtenido como se observa en la Tabla 8 y en la Figura 11.

#### Análisis y discusión comparativa

#### Análisis de pruebas mediante las matrices de confusión

En la Prueba 1 (YOLOv8n) se evidenció un bajo desempeño en

la detección de armas y personas armadas, con numerosos falsos negativos y confusiones hacia la clase background. Esto reflejó las limitaciones del modelo ligero para escenarios complejos. En la Prueba 2 (YOLOv8s) se registraron mejoras en la detección de personas, aunque persistieron falsos positivos entre arma y persona armada. En la Prueba 3 (YOLOv8s) se alcanzó un balance más favorable, reduciendo los errores en la clase persona, aunque la identificación de personas armadas continuó siendo un reto. Finalmente, la Prueba 4 (YOLOv8s) mostró el mejor desempeño, con una reducción significativa de falsos positivos y falsos negativos, logrando alta precisión en todas las clases críticas.



**Tabla 8.** Comparación entre valores previstos y reales por escenario.

Fuente: Los autores.

| Escenario (Video)                                  | Métrica   | Previsto (%) | Obtenido (%) | Diferencia |
|--|-----------|--------------|--------------|------------|
| Video 1 – Robo interior iluminado                  | Precisión | 95           | 92           | -3.0       |
|  | Recall    | 93           | 90           | -3.0       |
|  | F1-Score  | 94           | 91           | -3.0       |
|  | mAP@50    | 97           | 94           | -3.0       |
|  | mAP@50–95 | 76           | 72           | -4.0       |
| Video 2 – Tienda Interior iluminado                | Precisión | 95           | 93           | -2.0       |
|  | Recall    | 93           | 92           | -1.0       |
|  | F1-Score  | 94           | 92           | -2.0       |
|  | mAP@50    | 97           | 95           | -2.0       |
|  | mAP@50–95 | 76           | 74           | -2.0       |
| Video 3 – Robo local interior                      | Precisión | 95           | 93           | -2.0       |
|  | Recall    | 93           | 92           | -1.0       |
|  | F1-Score  | 94           | 92           | -2.0       |
|  | mAP@50    | 97           | 95           | -2.0       |
|  | mAP@50–95 | 76           | 73           | -3.0       |
| Video 4 – Escenario nocturno mixto                 | Precisión | 95           | 90           | -5.0       |
|  | Recall    | 93           | 88           | -5.0       |
|  | F1-Score  | 94           | 89           | -5.0       |
|  | mAP@50    | 97           | 92           | -5.0       |
|  | mAP@50–95 | 76           | 70           | -6.0       |
| Video 5 – Robo interior Medio iluminado            | Precisión | 95           | 91           | -4.0       |
|  | Recall    | 93           | 90           | -3.0       |
|  | F1-Score  | 94           | 90           | -4.0       |
|  | mAP@50    | 97           | 93           | -4.0       |
|  | mAP@50–95 | 76           | 71           | -5.0       |
| Video 6 – Actividades sospechosas                  | Precisión | 95           | 91           | -4.0       |
|  | Recall    | 93           | 90           | -3.0       |
|  | F1-Score  | 94           | 90           | -4.0       |
|  | mAP@50    | 97           | 93           | -4.0       |
|  | mAP@50–95 | 76           | 71           | -5.0       |
| Video 7 – Exterior con lluvia leve                 | Precisión | 95           | 92           | -3.0       |
|  | Recall    | 93           | 91           | -2.0       |
|  | F1-Score  | 94           | 92           | -2.0       |
|  | mAP@50    | 97           | 94           | -3.0       |
|  | mAP@50–95 | 76           | 73           | -3.0       |
| Video 8 – Interior con movimiento rápido de cámara | Precisión | 95           | 93           | -2.0       |
|  | Recall    | 93           | 91           | -2.0       |
|  | F1-Score  | 94           | 92           | -2.0       |
|  | mAP@50    | 97           | 94           | -3.0       |
|  | mAP@50–95 | 76           | 73           | -3.0       |
| Video 9 – Multitud en espacio público diurno       | Precisión | 95           | 92           | -3.0       |
|  | Recall    | 93           | 90           | -3.0       |
|  | F1-Score  | 94           | 91           | -3.0       |
|  | mAP@50    | 97           | 94           | -3.0       |
|  | mAP@50–95 | 76           | 72           | -4.0       |
| Video 10 – Interior nocturno con luz artificial    | Precisión | 95           | 91           | -4.0       |
|  | Recall    | 93           | 89           | -4.0       |
|  | F1-Score  | 94           | 90           | -4.0       |
|  | mAP@50    | 97           | 93           | -4.0       |
|  | mAP@50–95 | 76           | 71           | -5.0       |

## Análisis de la Matriz de confusión del mejor modelo

La Tabla 9 presenta la matriz de confusión correspondiente a las pruebas realizadas con el modelo seleccionado, donde se resumen las predicciones correctas e incorrectas para cada clase definida en el sistema: persona, arma y persona armada.

### El análisis reveló que:

- La mayor parte de los errores se concentraron en la confusión entre las clases arma y persona armada, ya que en varias ocasiones el modelo clasificó incorrectamente armas como personas armadas o viceversa. Esta situación puede atribuirse a la similitud visual y a la dificultad de identificar con precisión objetos pequeños en escenas con variaciones de ángulo o con occlusiones parciales.
- También se registraron casos en los que personas armadas fueron clasificadas únicamente como persona, lo que indica que la detección conjunta del individuo y el arma continúa siendo un reto en escenarios de mayor complejidad.
- En menor medida, se observaron confusiones de armas con objetos no etiquetados (como herramientas o teléfonos móviles), lo que sugiere la necesidad de refinar el dataset con más ejemplos negativos y diversificar las condiciones de entrenamiento para reducir estos falsos positivos.

Con base en los resultados obtenidos en la Tabla 6, la Prueba 4 (YOLOv8s) fue seleccionada como el mejor modelo debido a su rendimiento superior y más equilibrado en comparación con las configuraciones anteriores. Mientras que las pruebas iniciales evidenciaron limitaciones notables en la detección de armas y personas armadas, así como altos niveles de falsos positivos y negativos, la versión optimizada logró superar estas deficiencias alcanzando métricas de precisión (95.08 %), recall (92.94 %) y F1-score (93.99 %) que confirman su solidez en la clasificación de las tres clases críticas. En particular, el valor de mAP@50 (96.97 %) resalta su capacidad para identificar objetos con alta exactitud en condiciones diversas, mientras que el desempeño en mAP@50–95 (72.68 %) demuestra una robustez aceptable frente a variaciones en la superposición entre predicciones y etiquetas reales.

### Análisis comparativos del prototipo en diversos escenarios

- **Video 1 (Robo interior iluminado):** mantuvo cifras elevadas y similares a lo anticipado, con variaciones de hasta el 3 % en Recall y Precisión. La detección ininterrumpida de armas y personas fue posible gracias a la iluminación adecuada en un ambiente regulado, lo que asegura un desempeño confiable.
- **Video 2 (Tienda interior iluminado):** el modelo se demostró

eficaz, con cifras estables y una discrepancia de menos del 2 % respecto a lo que se esperaba. Se disminuyó el número de falsos negativos y la precisión se mantuvo en niveles altos debido a una iluminación homogénea, lo que permitió identificar mejor a las personas y a las armas.

- **Video 3 (Robo local interior):** este fue uno de los contextos más estables, con pérdidas mínimas y métricas que superan el 92 %. Un rendimiento parejo en la detección de todas las clases fue facilitado por un entorno cerrado y con condiciones homogéneas.

- **Video 4 (Escenario mixto):** mostró reducciones de hasta un 6 % en Recall y mAP@50-95, lo cual demuestra que el modelo es sensible a las condiciones difíciles de luz. La existencia de contrastes y sombras produjo ciertos falsos negativos, particularmente en el reconocimiento de armas.

- **Video 5 (Robo Interior medio iluminado):** exhibió un rendimiento aceptable, con descensos de entre el 3 % y el 5 %. El sistema mantuvo la estabilidad en la detección de individuos y una calidad aceptable en la identificación de armas, a pesar de que la iluminación no fue tan buena como en los primeros escenarios.

- **Video 6 (Actividad sospechosa):** mantuvo las métricas generales por encima del 90%, lo que confirma que el prototipo es capaz de funcionar en situaciones críticas. La identificación simultánea de personas y armas se llevó de manera confiable, aunque es importante señalar que sí hubo pérdidas mínimas en Recall, principalmente debido a la diversidad de situaciones en el escenario.

- **Video 7 – (Exterior con lluvia leve):** se analizó el desempeño en condiciones climáticas desfavorables. A pesar de que se notó un pequeño incremento en los falsos negativos de armas debido al desenfoque provocado por la lluvia, el modelo mantuvo detecciones estables.

- **Video 8 – (Interior con movimiento rápido de cámara):** se puso a prueba la resistencia del sistema frente a cambios súbitos de enfoque. Los resultados revelaron una buena estabilidad, manteniendo una alta exactitud en la identificación de personas y personas armadas.

- **Video 9 - (Multitud en espacio público diurno):** se estudió la capacidad de detección en un entorno con una gran cantidad de personas y objetos. A pesar de que se dieron confusiones específicas entre armas y objetos semejantes, el sistema logró sostener un equilibrio razonable.

- **Video 10 – (Interior nocturno con luz artificial):** se constató la eficacia en condiciones de iluminación mixta, mostrando una ligera reducción del recall frente a los entornos diurnos con luz natural; no obstante, se detectaron correctamente las tres clases críticas.

El análisis comparativo entre el peor y el mejor escenario analizado mostró de manera directa cómo las condiciones del entorno y la luz influyen en el desempeño del modelo. En

la Figura 12 (Video 1 -interior iluminado), se evidenció una detección confiable y estable de personas armadas, armas y personas, con métricas de precisión y recall que superaron el 90 %. La identificación precisa de los objetos fue posible gracias a la ausencia de occlusiones, al fondo uniforme y a la buena iluminación, lo que demostró que el modelo YOLOv8s puede funcionar con eficacia en escenarios controlados. Por otro lado, La Figura 13 (Video 4 - interior/exterior nocturno) fue el más difícil de todos, con caídas registradas en las métricas de recall y mAP@50-95 que alcanzaron hasta un 6 %. Las variaciones de luz, las sombras y los reflejos produjeron falsos negativos, sobre todo en la detección de armas, lo que demuestra que el modelo es sensible a condiciones desfavorables de iluminación.

## Discusión

El análisis de los resultados hizo posible el reconocimiento preciso de las fortalezas y limitaciones del prototipo sugerido, así como su contribución distintiva en el ámbito de la videovigilancia inteligente. El modelo YOLOv8s mostró un rendimiento sólido, conservando indicadores de precisión y recall superiores al 90% bajo condiciones óptimas de iluminación, lo cual confirma su utilidad para usos de videovigilancia en tiempo real. Este comportamiento demuestra que el modelo tiene la habilidad de equilibrar la velocidad de inferencia con la precisión en la identificación de personas y armas, como se ha informado en investigaciones recientes sobre la misma arquitectura (Delwar et al., 2025; Nasir et al., 2025). Sin embargo, se notó que el rendimiento disminuyó hasta un 6% en condiciones de poca luminosidad, lo cual demuestra que la calidad del entorno visual continúa siendo un factor clave para percibir objetos pequeños o parcialmente visibles.

El prototipo desarrollado cumple con el objetivo de la investigación, logrando un sólido F1-score del 93.99%. La eficiencia computacional fue otro elemento significativo que se examinó, además de la precisión. Este estudio, a diferencia de la mayoría de los trabajos existentes que se enfocan en optimizar las métricas de precisión, también verificó el desempeño del prototipo en hardware asequible. Según lo indicado en la Fase 6, el sistema final funciona con un tiempo total de procesamiento de 30 ms, que incluye 22 ms para la inferencia por cuadro. Este descubrimiento es importante cuando se comparan nuestros resultados con los de la literatura. Aunque los trabajos de Nasir et al. (2025) y Delwar et al. (2025) ofrecen F1-scores teóricos más altos, del 94.7% y del 98% respectivamente, suelen enfocarse solo en la precisión algorítmica sin tratar el costo computacional o las exigencias de hardware. Aunque un modelo que necesita una GPU de gama alta puede parecer atractivo, no es viable implementarlo de manera masiva. En contraste, nuestra investigación muestra que es factible lograr una precisión elevada con un F1-score de 93.99%, y a la vez ser prácticamente viable, funcionando de manera eficaz en hardware de gama baja y para el consumo.

Esta democratización de la tecnología tiene efectos directos en situaciones latinoamericanas y en países en vías de desarrollo



con recursos escasos. La arquitectura propuesta posibilita que otras instituciones apliquen el sistema sin depender de soluciones costosas y de propiedad exclusiva, al facilitar su replicabilidad por medio de herramientas de código abierto. A diferencia de esto, los estudios anteriores, como Schcolnik-Elias et al. (2023) y Gawande et al. (2024), optimizan la precisión, pero no toman en cuenta la escalabilidad económica. Los hallazgos evidencian que este enfoque no pone en riesgo la eficacia del sistema. Cuando se pondera la disminución de costos, la diferencia en rendimiento con respecto al hardware de gama alta es mínima; este hallazgo es especialmente significativo en un contexto donde la brecha tecnológica en aplicaciones de inteligencia artificial para la seguridad pública sigue creciendo.

Los hallazgos, en esencia, confirman la eficacia del prototipo y su aptitud para incorporarse a sistemas reales de vigilancia inteligente que cuenten con hardware comercial accesible. El hecho de que la calidad del video y la luminosidad ambiental dependan mutuamente es una limitación operativa intrínseca a los sistemas de visión por computadora. Para solucionar este problema, se deben utilizar en etapas posteriores sensores infrarrojos o métodos avanzados de aumento de datos específicos para condiciones con poca luz. Estas mejoras no solamente aumentarían la robustez y adaptabilidad del sistema en situaciones adversas, sino que también fortalecerían su aportación como una herramienta tecnológica escalable, factible y asequible para la seguridad ciudadana actual. Esto es particularmente importante en escenarios con recursos escasos en donde la necesidad de sistemas inteligentes para prevenir y responder rápidamente a amenazas es cada vez más urgente.

#### 4. Conclusiones

El estudio fue capaz de crear y validar un prototipo de monitoreo inteligente fundamentado en YOLOv8s, el cual muestra un balance ideal entre una precisión alta y la eficiencia computacional. Los resultados corroboran que el modelo es una herramienta sólida para la detección en tiempo real, lo cual valida tanto su factibilidad técnica como su viabilidad económica al funcionar de manera eficaz en hardware asequible. Por ende, este trabajo proporciona una solución práctica y con capacidad de escalarse económicamente, que cierra la brecha entre la exactitud teórica de los modelos de IA y su implementación posible en sistemas de seguridad reales. La metodología CRISP-DM permitió la adecuada organización de las fases de modelado, evaluación y preparación, garantizando de esta manera un proceso laboral que es eficiente y reproducible.

Sin embargo, el estudio también estableció limitaciones

inherentes al diseño de la investigación y a la naturaleza del conjunto de datos. La evaluación del desempeño en contextos no estructurados o con cambios drásticos está restringida por la dependencia de videos breves y regulados. Asimismo, la dificultad para detectar objetos y armas no letales y la sensibilidad del modelo en condiciones de escasa iluminación muestran que su exactitud de detección aún se encuentra sujeta a factores ajenos. A pesar de que estos límites no invalidan los descubrimientos, resaltan la importancia de fortalecer la robustez del modelo utilizando un conjunto de datos más diverso y representativo.

A partir de estas observaciones, se proponen múltiples líneas de investigación futura para optimizar el sistema: aumentar la diversidad del conjunto de datos incorporando escenas más diversas, implementar técnicas de aumento de datos más avanzadas, introducir sensores térmicos o infrarrojos en situaciones con poca luminosidad y analizar arquitecturas híbridas de detección que integren diferentes clases de redes neuronales. Estos progresos tienen el potencial de reducir los falsos positivos que surgen al clasificar objetos y mejorar la habilidad del sistema para operar en condiciones adversas.

#### Contribución de los autores

**Lauro Alfonso Erreyes Cuenca:** Administración del proyecto, investigación, redacción y metodología. **Nahin Josue Olmedo Chica:** Investigación, y edición del artículo. **Zea Mariuxi:** Metodología, revisión, redacción y edición del artículo. **Nancy Magaly Loja Mora:** Metodología, revisión, redacción y edición del artículo.

#### Conflictos de interés

Los autores no tienen conflictos de interés.

#### Referencias bibliográficas

- Acuña-Cid, H. A., Ahumada-Tello, E., Ovalle-Osuna, Ó. O., Evans, R., Hernández-Ríos, J. E., & Zambrano-Soto, M. A. (2025). CRISP-NET: Integration of the CRISP-DM Model with Network Analysis. *Machine Learning and Knowledge Extraction*, 7(3), 101. <https://doi.org/10.3390/make7030101>
- Alvarado, H. P. L., & Villavicencio, O. E. C. (2024). Regulación del Manejo de la Inteligencia Artificial, Consecuencias y Daños a la Sociedad por su Mal Uso. *Ciencia Latina Revista Científica Multidisciplinar*, 8(1), 1966–1978. [https://doi.org/10.37811/cl\\_rcm.v8i1.9596](https://doi.org/10.37811/cl_rcm.v8i1.9596)

- Azatbekuly, N., Mukhanbet, A., & Bekele, S. D. (2024). Development of an Intelligent Video Surveillance System Based on YOLO Algorithm. 2024 IEEE 4th International Conference on Smart Information Systems and Technologies (SIST), 498–503. <https://doi.org/10.1109/SIST61555.2024.10629617>
- Castro-Paredes, J. G., Mendoza-Masache, G. R., Loja-Mora, N. M., & Loján-Alvarado, H. P. (2025). Protección de datos por diseño y por defecto. Implicaciones legales en el desarrollo de software. *Ingenium et Potentia*, 7(12), 77–96. <https://doi.org/10.35381/i.p.v7i12.4471>
- Cheng, G., Chao, P., Yang, J., & Ding, H. (2024). SGST-YOLOv8: An Improved Lightweight YOLOv8 for Real-Time Target Detection for Campus Surveillance. *Applied Sciences*, 14(12), 5341. <https://doi.org/10.3390/app14125341>
- Delwar, T. S., Mukhopadhyay, S., Kumar, A., Singh, M., Lee, Y., Ryu, J.-Y., & Hosen, A. S. M. S. (2025). Real-Time Farm Surveillance Using IoT and YOLOv8 for Animal Intrusion Detection. *Future Internet*, 17(2), 70. <https://doi.org/10.3390/fi17020070>
- Erreyes, L. (2025). Dataset Detección personas-armas-personas armadas. Zenodo. <https://doi.org/10.5281/zenodo.17215185>
- Espinoza, L. J. M., Rojas, S. J. Z., Bravo, V. J. P., & Valarezo, L. C. C. (2025). Balanceo de Conjuntos de Datos Basado en Redes Generativas Aplicado a Imágenes del Sector Agrícola. *Informática y Sistemas*, 9(2), 164–176. <https://doi.org/10.33936/isrtic.v9i2.7782>
- Gawande, U., Hajari, K., & Golhar, Y. (2024). Novel person detection and suspicious activity recognition using enhanced YOLOv5 and motion feature map. *Artificial Intelligence Review*, 57(2), 16. <https://doi.org/10.1007/s10462-023-10630-0>
- Hua, C., Luo, K., Wu, Y., & Shi, R. (2024). YOLO-ABD: A Multi-Scale Detection Model for Pedestrian Anomaly Behavior Detection. *Symmetry*, 16(8), 1003. <https://doi.org/10.3390/sym16081003>
- Li, Y., Li, Q., Pan, J., Zhou, Y., Zhu, H., Wei, H., & Liu, C. (2024). SOD-YOLO: Small-Object-Detection Algorithm Based on Improved YOLOv8 for UAV Images. *Remote Sensing*, 16(16), 3057. <https://doi.org/10.3390/rs16163057>
- Nasir, R., Jalil, Z., Nasir, M., Alsubait, T., Ashraf, M., & Saleem, S. (2025). An enhanced framework for real-time dense crowd abnormal behavior detection using YOLOv8. *Artificial Intelligence Review*, 58(7), 202. <https://doi.org/10.1007/s10462-025-11206-w>
- Ni, Y., Huo, J., Hou, Y., Wang, J., & Guo, P. (2024). Detection of Underground Dangerous Area Based on Improving YOLOV8. *Electronics*, 13(3), 623. <https://doi.org/10.3390/electronics13030623>
- Schcolnik-Elias, A., Martínez-Díaz, S., Luna-Taylor, J. E., & Castro-Liera, I. (2023). Detección de armas tipo pistola mediante el uso de redes convolucionales con una arquitectura tipo YOLO y estereoscopía. *Pádi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI*, 11, 196–204. <https://doi.org/10.29057/icbi.v11iEspecial2.10727>
- Stainer, M. J., Raj, P. V., Aitken, B. M., Bandarian-Balooch, S., & Boschen, M. J. (2021). Decision-making in single and multiple-screen CCTV surveillance. *Applied Ergonomics*, 93, 103383. <https://doi.org/10.1016/j.apergo.2021.103383>
- Sudharson, D., Srinithi, J., Akshara, S., Abhirami, K., Sriharshitha, P., & Priyanka, K. (2023). Proactive Headcount and Suspicious Activity Detection using YOLOv8. *Procedia Computer Science*, 230, 61–69. <https://doi.org/10.1016/j.procs.2023.12.061>

