



Facultad de Ciencias Sociales, Derecho y Bienestar

Carrera de Comunicación

**MODALIDAD: ESCRITURA DE ARTÍCULO CIENTÍFICO O
CAPÍTULO DE LIBRO**

**Previo a la obtención del título de grado de:
LICENCIADO(A) EN COMUNICACIÓN**

TEMA:

**HUMANOS VS MÁQUINAS: UN ESTUDIO COMPARATIVO
SOBRE LA HABILIDAD HUMANA PARA DETECTAR CONTENIDO
GENERADO POR IA**

LÍNEA DE INVESTIGACIÓN:

Comunicación basada en datos y tecnologías emergentes

AUTOR:

Miguel Angel Zambrano Romero

TUTOR:

Dr. Julio Cesar García García, PhD.

2026

TEMA:

**HUMANOS VS MÁQUINAS: UN ESTUDIO COMPARATIVO SOBRE
LA HABILIDAD HUMANA PARA DETECTAR CONTENIDO GENERADO
POR IA**

DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHO DEL AUTOR

Yo, MIGUEL ANGEL ZAMBRANO ROMERO, portador de la cédula de ciudadanía No. 1314998905, declaro que el presente trabajo de investigación, y criterios emitidos, respetan rigurosamente en todo momento las normas éticas, previstas en la Ley de Propiedad Intelectual, asumo responsabilidad en lo referente a criterios, doctrinas, que contenga el trabajo de investigación, titulado: "HUMANOS VS MÁQUINAS: UN ESTUDIO COMPARATIVO SOBRE LA HABILIDAD HUMANA PARA DETECTAR CONTENIDO GENERADO POR IA", son de mi autoría, y autorizo a la Universidad Laica Eloy Alfaro de Manabí, para que haga de la investigación o parte de ella, documento disponible para, consultas de investigación, según las normas de la Institución, cediendo y aprobando la reproducción de los derechos patrimoniales del trabajo, con fines de difusión pública, dentro de las regulaciones de la Universidad, siempre y cuando esta reproducción, no suponga ganancia económica, y realice los derechos que como autor me corresponden, con excepción de la presente autorización, seguirán vigentes a mi favor, de conformidad con lo establecido en los artículos 5, 6, 8, 19 de la Ley de Propiedad Intelectual y su Reglamento, en concordancia, con el Art. 144 de la Ley de Educación Superior.

Para constancia firmo a continuación.

Miguel Zambrano.

Miguel Angel Zambrano Romero
C.I. 1314998905

CERTIFICADO DEL TUTOR

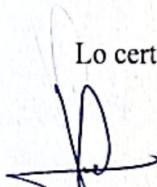
En calidad de docente tutor(a) de la carrera de **Comunicación** de la **Facultad de Ciencias Sociales, Derecho y Bienestar** de la Universidad Laica "Eloy Alfaro" de Manabí, certifico:

Haber dirigido y revisado el trabajo de titulación del alumno MIGUEL ANGEL ZAMBRANO ROMERO, estudiante de la carrera de Comunicación, período académico 2026(1); cumpliendo el total de 400 horas, bajo la modalidad de **Escritura de Artículo Científico o Capítulo de Libro**, cuyo tema es "HUMANOS VS MÁQUINAS: UN ESTUDIO COMPARATIVO SOBRE LA HABILIDAD HUMANA PARA DETECTAR CONTENIDO GENERADO POR IA". Dicho trabajo ha sido desarrollado de acuerdo con los lineamientos internos de la modalidad en mención y en apego al cumplimiento de los requisitos exigidos por el Reglamento de Régimen Académico, razón por la cual reúne los méritos académicos, científicos y formales suficientes para ser sometido a la evaluación del tribunal de titulación que designe la autoridad competente. En consecuencia, el(la) estudiante en mención se encuentra apto(a) para la sustentación de su trabajo de titulación.

Particular que certifico para los fines consiguientes, salvo disposición de Ley en contrario.

Manta, 18 de agosto de 2026

Lo certifico,



Dr. Julio César García García, PhD.
DOCENTE TUTOR

AGRADECIMIENTO

Agradezco primeramente a Dios, por darme la salud, fortaleza y sabiduría necesarias para culminar esta etapa de mi formación profesional.

A mis padres, por su apoyo incondicional y por brindarme siempre las herramientas necesarias para alcanzar mis metas.

A la Universidad Laica Eloy Alfaro de Manabí, por abrirme sus puertas y darme la oportunidad de formarme profesionalmente en sus aulas.

A mi tutor, el Dr. Julio César García, por su guía, paciencia y valiosos aportes durante el desarrollo de este trabajo de titulación.

A los docentes que a lo largo de mi carrera compartieron sus conocimientos y experiencia, contribuyendo a mi formación como profesional.

DEDICATORIA

A mis padres, Miguel Zambrano y Flor Romero, por su amor, esfuerzo y sacrificio para darme la oportunidad de estudiar y lograr esta meta académica.

A mi hermana por su compañía en cada etapa de este camino.

A mis abuelos maternos por su cariño y por estar presentes en cada paso de mi vida.

A mi novia, Iris, por su paciencia, motivación y por creer en mí incluso en los momentos más difíciles.

CAPTURA ENVÍO ARTÍCULO

[EAC] Acuse de recibo del envío  Resumir este correo electrónico



Rita Milagros Jáimez Esteves <revistas.noreply@unl.edu.ec>

Para: GARCIA GARCIA JULIO CESAR



Mar 4/8/2025 21:04

Dr. Julio García, PhD.:

Gracias por enviar el manuscrito "Humanos vs máquinas: un estudio comparativo sobre la habilidad humana para detectar contenido generado por IA" a Educación, Arte, Comunicación: Revista Académica e Investigativa. Con el sistema de gestión de publicaciones en línea que utilizamos podrá seguir el progreso a través del proceso editorial tras iniciar sesión en el sitio web de la publicación:

URL del manuscrito: <https://revistas.unl.edu.ec/index.php/eac/authorDashboard/submission/2850>

Nombre de usuario/a: jcgarcia1805

Si tiene alguna duda puede ponerse en contacto conmigo. Gracias por elegir esta editorial para mostrar su trabajo.

Rita Milagros Jáimez Esteves

Educación Arte Comunicación

<http://revistas.unl.edu.ec/index.php/eac>

Humanos vs máquinas: un estudio comparativo sobre la habilidad humana para detectar contenido generado por IA

Autor: Miguel Angel Zambrano Romero
ORCID: <https://orcid.org/0009-0006-9488-9542>
zambranomiguel3018@gmail.com
Telf. +593 961358473
Código postal: 130221
Universidad Laica Eloy Alfaro de Manabí

Dr. Julio César García García, PhD.
ORCID: <https://orcid.org/0000-0002-3381-1147>
Universidad Laica Eloy Alfaro de Manabí
julio.garcia@uleam.edu.ec
Telf. +593 984677610
Código postal: 130221
Universidad Laica Eloy Alfaro de Manabí

Resumen

El rápido avance de la inteligencia artificial (IA) plantea nuevos desafíos para la percepción humana y la validación de la información. El objetivo de este estudio fue determinar la habilidad de los estudiantes de la ULEAM para reconocer entre imágenes de autoría humana e imágenes generadas por inteligencia artificial.

Mediante un diseño cuantitativo basado en la Teoría de Detección de Señales, se aplicó un instrumento de evaluación visual de forma presencial en las aulas a una muestra de 381 estudiantes. Los estímulos incluyeron imágenes reales, así como sintéticas generadas mediante modelos de IA generativa. Los resultados mostraron una vulnerabilidad perceptiva generalizada, con un índice de detectabilidad débil-moderado ($d' = 0.63$) y un sesgo de respuesta conservador ($c = 0.61$), señalando una predisposición a asumir las imágenes sintéticas como auténticas. No se hallaron diferencias significativas según el género o experiencia previa con IA. Sin embargo, se registró un rendimiento superior en las facultades de Ciencias Sociales ($d' = 0.81$)

y Ciencias de la Salud ($d' = 0.69$), con una discriminación moderada en los estudiantes de noveno semestre ($d' = 1.01$), registrando un mayor rendimiento en estudiantes de niveles superiores, lo que sugiere una asociación positiva entre el avance académico y el escrutinio crítico. El análisis de indicios visuales demostró que la detección exitosa requiere abandonar la intuición superficial para aplicar un escrutinio técnico enfocado en el detalle y el color. Se concluye que la intuición visual es insuficiente frente a la IA generativa, exigiendo una alfabetización mediática estructural y transversal para reducir el engaño cognitivo.

Palabras Clave: alfabetización mediática, fotorrealismo, inteligencia artificial, percepción visual, Teoría de Detección de Señales.

Abstract

The rapid advancement of artificial intelligence (AI) poses new challenges for human perception and information validation. The objective of this study was to determine the ability of ULEAM students to distinguish between human-authored images and images generated by artificial intelligence. Through a quantitative design based on Signal Detection Theory, a visual evaluation instrument was administered in-person within the classrooms to a sample of 381 students. The stimuli included real images, as well as synthetic ones generated through generative AI models. The results showed a generalized perceptual vulnerability, with a weak-to-moderate detectability index ($d' = 0.63$) and a conservative response bias ($c = 0.61$), indicating a predisposition to assume synthetic images as authentic. No significant differences were found based on gender or prior experience with AI. However, superior performance was recorded in the faculties of Social Sciences ($d' = 0.81$) and Health Sciences ($d' = 0.69$), with a moderate discrimination value among ninth-semester

students ($d' = 1.01$), recording higher performance in upper-level students, suggesting a positive association between academic advancement and critical scrutiny. The analysis of visual cues demonstrated that successful detection requires abandoning superficial intuition to apply technical scrutiny focused on detail and color. It is concluded that visual intuition is insufficient against generative AI, demanding structural and transversal media literacy to reduce cognitive deception.

Keywords: artificial intelligence, media literacy, photorealism, Signal Detection Theory, visual perception.

Introducción

El rápido avance de la inteligencia artificial se ha consolidado como una tecnología disruptiva que está transformando radicalmente los ecosistemas de información a escala global. Su capacidad para producir contenido sintético textual, visual y audiovisual con un nivel de coherencia y naturalidad cada vez más indistinguible del creado por seres humanos representa un punto de inflexión sin precedentes en la historia de la comunicación (World Economic Forum, 2025). Este fenómeno no solo amplifica el volumen y la velocidad de la desinformación, sino que eleva su credibilidad percibida hasta niveles que dificultan gravemente su detección, con repercusiones directas en áreas críticas como la salud pública (Saeidnia et al., 2026).

La evidencia empírica sobre la vulnerabilidad humana ante este escenario es contundente y, al mismo tiempo, preocupante. Un reciente estudio determinó que la precisión promedio de las personas para detectar deepfakes y contenido sintético de alta calidad es de apenas el 55,54%, un desempeño que no supera estadísticamente el nivel del azar (Diel et al., 2024). Por otro lado, Pennycook y Rand (2021) ofrecen un sustento cognitivo para comprender la vulnerabilidad de las personas frente a la desinformación mediante el modelo de procesamiento dual. Según los autores, el Sistema 1 opera de manera rápida, automática e intuitiva, mientras que el Sistema 2 se caracteriza por un procesamiento más lento, deliberativo y analítico. En este sentido, cuando los individuos se enfrentan a información falsa que resulta fluida, coherente o aparentemente creíble, pueden apoyarse principalmente en las respuestas intuitivas del Sistema 1 actuando como lo que los autores denominan "avaros cognitivos" sin activar suficientemente los procesos reflexivos del Sistema 2, indispensables para evaluar críticamente la veracidad de los contenidos y detectar

posibles formas de engaño algorítmico. Así, el desarrollo técnico de la IA generativa explota precisamente las limitaciones estructurales del procesamiento cognitivo humano, lo que convierte a esta vulnerabilidad no en una particularidad individual, sino en un patrón sistemático y predecible.

Este problema global toma una dimensión crítica en América Latina, región donde la manipulación de información es reconocida como una amenaza. De acuerdo con datos de Activa Research y WIN (2023), el 73% de la población latinoamericana percibe la desinformación en las noticias como un problema importante. Este porcentaje se incrementa al 78% específicamente en Ecuador. Esta preocupación trasciende el plano informativo, ya que el 77% de los latinoamericanos considera que la desinformación debilita los procesos electorales y un 80% la percibe como una amenaza directa para la democracia. Estos riesgos se intensifican en el contexto ecuatoriano con el aumento de las deepfakes, las cuales debilitan la confianza pública y la cohesión social (Altamirano Molina, 2026).

Dentro de este escenario regional, los universitarios aparecen como un grupo especialmente vulnerable. Investigaciones recientes evidencian que estudiantes del área de comunicación no solo son incapaces de distinguir entre noticias falsas generadas por IA y noticias verdaderas redactadas por humanos, sino que, paradójicamente, llegan a percibir el contenido sintético como más veraz, preciso y claro que la información periodística auténtica (García-Marín et al., 2025). Este hallazgo resulta particularmente revelador porque contradice la suposición común de que la formación universitaria, y especialmente aquella vinculada al campo de la comunicación, debería desempeñarse como un factor protector frente a la desinformación. Si esta vulnerabilidad se manifiesta ante textos escritos que permiten una lectura reflexiva, el riesgo adquiere dimensiones más críticas frente a

imágenes fotorrealistas. La evidencia empírica reciente demuestra que el escrutinio de imágenes estáticas generadas por inteligencia artificial provoca sesgos cognitivos sistemáticos que alteran la evaluación analítica rigurosa, forzando a los individuos a operar bajo juicios heurísticos que mantienen la tasa de acierto humana en niveles estadísticamente equivalentes al azar (Casu et al., 2025). En este sentido, los datos sugieren que la exposición formativa sin alfabetización mediática estrictamente visual resulta insuficiente ante el avance de las herramientas generativas.

En el contexto ecuatoriano, las consecuencias de esta vulnerabilidad ya han generado repercusiones sociales graves y documentadas. En octubre de 2023, estudiantes menores de edad en una institución educativa de Quito utilizaron IA para crear contenido sexual falso de sus compañeras, un caso que expuso profundos vacíos normativos, éticos y legales en la regulación nacional frente a estas tecnologías (Altamirano Molina, 2026). A nivel universitario, un estudio llevado a cabo con estudiantes de ingeniería en la Escuela Superior Politécnica del Litoral (ESPOL) confirmó empíricamente esta incapacidad, al demostrar que no existe una diferencia estadísticamente significativa en la credibilidad que los universitarios otorgan a noticias reales frente a las hechas con IA (El Salous et al., 2026). Este resultado, leído en conjunto con los hallazgos de García-Marín et al. (2025), permite establecer un patrón generalizado que va más allá de la disciplina o del perfil formativo del estudiante, reflejado en la incapacidad para detectar textos y noticias generadas por IA. Ante el vacío en la literatura sobre cómo opera esta vulnerabilidad frente a estímulos puramente visuales, la incapacidad para identificar imágenes fotorrealistas representa una vulnerabilidad emergente que requiere ser investigada empíricamente, agravada por una baja alfabetización mediática visual y un escaso

conocimiento sobre las implicaciones legales de compartir contenidos fabricados en plataformas digitales (Vega Cocha et al., 2024).

A pesar de la abrumadora evidencia sobre el impacto de la IA en la propagación de desinformación y la vulnerabilidad evidenciada en investigaciones recientes sobre jóvenes universitarios en Ecuador, persiste un notable vacío de conocimiento empírico en la provincia de Manabí. Hasta la fecha, la revisión de la literatura evidencia un vacío empírico respecto a la evaluación de la capacidad de detección de imágenes generadas por IA específicamente en los estudiantes de la Universidad Laica Eloy Alfaro de Manabí (ULEAM), ubicada en la ciudad de Manta, lo que impide dimensionar el alcance del problema en esta comunidad educativa y justifica plenamente la presente investigación.

A partir de este vacío, y considerando la literatura revisada, se formulan las siguientes hipótesis de investigación: **H1.** Existe una diferencia estadísticamente significativa en la habilidad para identificar imágenes generadas por IA entre los estudiantes con experiencia previa en su creación y aquellos sin experiencia. **H2.** Hay diferencias estadísticamente significativas en la habilidad para detectar imágenes creadas con IA según el género de los participantes. Para poner a prueba estas hipótesis y abordar el problema planteado, el presente estudio tiene como objetivo determinar la habilidad de los estudiantes de la Universidad Laica Eloy Alfaro de Manabí para distinguir entre imágenes generadas por inteligencia artificial e imágenes de autoría humana durante el año 2026. Para alcanzar este propósito, se establecen los siguientes objetivos específicos: comparar la habilidad de detección entre estudiantes con experiencia previa en la creación de imágenes mediante IA y aquellos sin esa experiencia; analizar las diferencias en la habilidad de detección según el género de los participantes; e identificar los indicios visuales que los

estudiantes utilizan con mayor frecuencia para justificar su reconocimiento de imágenes generadas por inteligencia artificial.

El artículo se organiza en seis secciones. Tras esta introducción, se presenta el marco referencial, donde se revisan los enfoques teóricos y antecedentes relevantes del estudio. Luego, se describe el diseño metodológico, detallando los procedimientos utilizados para la recolección y análisis de datos. Después, se exponen los resultados obtenidos y su discusión en relación con la literatura revisada, para finalmente presentar las conclusiones del estudio.

Marco referencial

Fotorrealismo y la ontología de la imagen sintética

La inteligencia artificial generativa crea contenido visual fotorrealista inédito mediante arquitecturas de *deep learning*, como los modelos de difusión y Redes Generativas Antagónicas (GAN) (Cruz-Meléndez, 2025; Nightingale & Farid, 2022; Sánchez Díaz, 2024). Sin embargo, a diferencia de la fotografía, que registra lumínicamente la realidad, las imágenes sintéticas son simulaciones algorítmicas predictivas sin anclaje físico con el entorno (Marín Álvarez et al., 2025). Esta desconexión ontológica no es solo una característica técnica, sino el origen de una nueva forma de engaño visual. Al fabricar evidencias visuales hiperrealistas que resultan casi indistinguibles del registro auténtico, esta ontología sintética desafía frontalmente las capacidades del observador humano y exige una reevaluación de cómo procesamos la información visual (Cruz-Meléndez, 2025; Marín Álvarez et al., 2025).

Engaño cognitivo y vulnerabilidad perceptiva

La dificultad humana para reconocer esta desconexión ontológica se explica a través de la Teoría del Procesamiento Dual (Pennycook & Rand, 2021). Cuando los individuos evalúan imágenes hiperrealistas, tienden a depender de la intuición y el procesamiento rápido (Sistema 1), asumiendo la información visual como verdadera por defecto, en lugar de activar un escrutinio analítico y deliberativo (Sistema 2). Para medir empíricamente esta vulnerabilidad y separar la sensibilidad perceptiva real de los sesgos de decisión del observador, resulta indispensable aplicar la Teoría de Detección de Señales (TDS) (Macmillan & Creelman, 2005). Los estudios previos reflejan un fallo sistemático en el Sistema 2: la tasa de error al clasificar estas imágenes promedia un 38,7% (Lu et al., 2023), y ante modelos recientes como FLUX.1-dev, la precisión cae al 29,04% (Högemann et al., 2025). Este fenómeno genera un exceso de confianza injustificado e induce sesgos donde los observadores perciben los rostros sintéticos como incluso más confiables que los reales (Nightingale & Farid, 2022), facilitando escenarios complejos de desinformación (Sánchez Díaz, 2024).

Alfabetización mediática frente a la desinformación

Para mitigar la vulnerabilidad cognitiva evidenciada a través de la Teoría de Detección de Señales, la intuición visual de los usuarios debe complementarse estrictamente con competencias críticas que fuercen la activación del Sistema 2 (Bray et al., 2023). En este contexto, la alfabetización mediática se posiciona como la respuesta estructural frente al avance de la ontología sintética. Dado que los jóvenes universitarios, por su alto consumo digital, son especialmente vulnerables a la desinformación, esta alfabetización debe trascender el uso funcional de

herramientas hacia un enfoque profundo (García-Delgado Giménez & Revilla Guijarro, 2025). Fomentar el pensamiento crítico resulta indispensable para formar ciudadanos capaces de deconstruir analíticamente los estímulos visuales y superar el engaño perceptivo en ecosistemas digitales saturados.

Metodología

Enfoque y tipo de investigación

El presente estudio se desarrolló bajo un enfoque cuantitativo, orientado a la medición empírica y al análisis estadístico de datos numéricos. La investigación es de alcance descriptivo-comparativo, ya que no busca establecer correlaciones entre variables, sino contrastar las diferencias perceptivas entre distintos grupos, con un diseño no experimental de corte transversal, dado que la recolección de la información se efectuó en un único momento temporal durante el año 2026.

Población y muestra

La población estuvo constituida por los estudiantes de la sede principal de la Universidad Laica Eloy Alfaro de Manabí (ULEAM), ubicada en la ciudad de Manta. Considerando que el último informe preliminar de rendición de cuentas publicado por la institución documenta una población activa de 17.474 estudiantes matriculados en el campus matriz durante el período académico 2024-2 (Universidad Laica Eloy Alfaro de Manabí [ULEAM], 2025), se utilizó esta cifra formal como línea base para calcular matemáticamente el tamaño muestral necesario y lograr la representatividad estadística en poblaciones finitas. Se aplicó la siguiente fórmula:

$$n = \frac{NZ^2pq}{e^2(N-1) + Z^2pq}$$

Donde N es el tamaño de la población (17.474), el nivel de confianza se estableció al 95% ($Z = 1.96$), asumiendo una varianza máxima ($p = 0.5$ y $q = 0.5$) y un margen de error del 5% ($e = 0.05$). Al resolver la ecuación, la muestra mínima requerida se mantuvo en 376 individuos.

Al haber validado exitosamente las respuestas de 381 estudiantes seleccionados mediante un muestreo probabilístico por conglomerados, el estudio superó la cantidad mínima requerida. El marco muestral se estructuró utilizando las aulas físicas como unidades primarias. El marco muestral estuvo constituido por todas las aulas de la universidad. El mecanismo de aleatorización consistió en sortear aulas al azar y evaluarlas consecutivamente hasta superar la muestra mínima, lo que resultó en 21 aulas visitadas. La selección se realizó aleatoriamente a nivel de aula; la participación individual fue voluntaria, por lo que la inferencia se interpreta considerando ambas etapas. La asimetría en la distribución final de facultades responde exclusivamente a la variabilidad de este proceso de azar. La participación fue estrictamente voluntaria, anónima y consolidada mediante un consentimiento informado. En cuanto a la operacionalización de las variables evaluadas para responder a todos los objetivos e hipótesis planteadas se detalla en la Matriz de Operacionalización (véase Tabla 1).

Tabla 1

Matriz de Operacionalización de Variables

Variable (Alineación)	Dimensión	Indicador	Ítem / Instrumento
V. Independiente: Tipo de estímulo visual (<i>Obj. General</i>)	Origen de la imagen	Fotorrealismo y ontología visual.	Estímulo dicotómico presentado en el cuestionario: 1. Fotografía

			real, 2. Imagen sintética.
V. Dependiente: Capacidad de detección (<i>Obj. General</i>)	Parámetros de TDS	Sensibilidad perceptiva (d') y Sesgo de respuesta (c). Exactitud global (ACC).	Clasificación del participante frente a 50 estímulos visuales (Matriz de Aciertos y Falsas Alarmas).
V. de Hipótesis 1: Experiencia previa con IA	Uso tecnológico	Nivel de experiencia práctica en generación de imágenes.	Pregunta cerrada: ¿Tiene experiencia previa creando imágenes con Inteligencia Artificial? (Sí / No).
V. de Hipótesis 2: Género	Identidad sociodemográfica	Femenino, Masculino.	Pregunta cerrada: ¿Cuál es su género?
V. de Análisis: Indicios visuales (<i>Obj. Específico 3</i>)	Criterios de evaluación visual	Frecuencia de selección de anomalías estructurales.	Pregunta de opción múltiple: Justifique su decisión (Detalle, Color, Suavidad, Intuición, etc.).
V. Complementaria 1: Facultad (<i>Análisis adicional</i>)	Formación académica	Facultades pertenecientes a la institución.	Pregunta de opción múltiple: ¿A qué facultad pertenece usted?
V. Complementaria 2: Semestre (<i>Análisis adicional</i>)	Avance académico	Ciclo de estudio actual (1ro al 10mo).	Pregunta de opción múltiple: ¿En qué semestre cursa actualmente?

Nota. TDS = Teoría de Detección de Señales; ACC = Exactitud global; d' = Índice de sensibilidad perceptiva; c = Sesgo de respuesta.

Técnicas e instrumentos de recolección de datos

La técnica empleada fue la encuesta, aplicada a través de un cuestionario interactivo en línea estructurado en la plataforma Google Forms. El instrumento es una adaptación técnica de Human Perception Benchmark de Lu et al. (2023). Previo al despliegue masivo, se ejecutó una prueba piloto con 28 estudiantes. Como resultados de este pilotaje, se definieron los ajustes metodológicos del reactivo: la descripción inicial de la investigación online se mantuvo breve, enfocada en la concientización, y se omitió específicamente cualquier mención sobre la medición de la capacidad humana para prevenir sesgos en los participantes. Por otro lado, la estandarización de los estímulos visuales se logró mediante la técnica de correspondencia estricta de prompts, asegurando que cada réplica sintética compartiera exactamente los mismos parámetros compositivos que su contraparte real. De manera independiente, el proceso de pilotaje y ajuste metodológico sirvió para afinar detalles del instrumento base a partir de los resultados de la prueba piloto. Finalmente, el instrumento completo empleado para el estudio se encuentra disponible para su verificación en el siguiente enlace externo: https://drive.google.com/file/d/1-yk_UkCmP_wVAFtk7GLBswmaNvimWGEz/view?usp=sharing

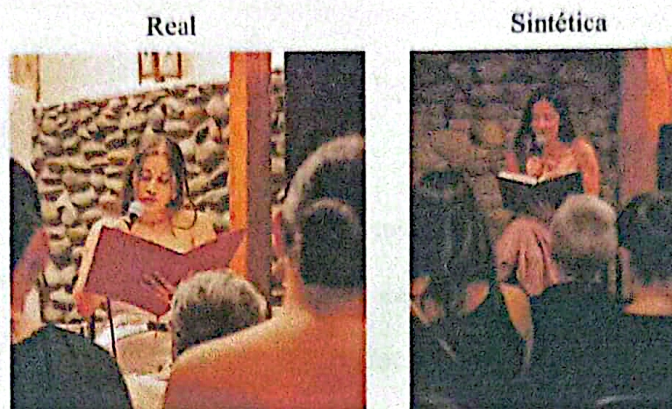
El cuestionario constó de 50 estímulos visuales balanceados, 25 imágenes reales y 25 imágenes sintéticas. Las fotografías auténticas fueron capturadas por el fotógrafo semiprofesional Joshue Delgado Pinto, estudiante de noveno semestre de la carrera de Comunicación de la ULEAM. Esto permitió documentar el procedimiento y los parámetros de iluminación, la originalidad del encuadre y la

ausencia de indexación previa en internet. Las imágenes sintéticas se generaron a través de la versión de pago de la plataforma de inteligencia artificial Magnific AI, utilizando los modelos algorítmicos integrados en su interfaz tales como: Google Nano Banana 2, GPT-2, Seedream 5 Lite y Flux.2 Pro. El uso de estas herramientas facilitó la obtención de estímulos visuales con un alto grado de fotorrealismo. Las versiones de generación corresponden a mayo de 2026, distribuyendo los 25 estímulos en: 6 con Flux.2 Pro, 12 con GPT-2, 3 con Seedream 5 Lite y 4 con Google Nano Banana 2, minimizando anomalías geométricas o artefactos evidentes para asegurar un nivel de complejidad elevado y estandarizado en la prueba de discriminación perceptiva. Cabe declarar como limitación que no se documentaron los *prompts* completos ni los parámetros exactos de generación empleados para dichos estímulos.

Para ilustrar el alto grado de fotorrealismo mencionado y el proceso de construcción del instrumento, se extrajo una descripción textual detallada (prompt) de fotografías auténticas de referencia para procesarlas mediante los modelos algorítmicos descritos. Este procedimiento garantizó que la réplica sintética compartiera las mismas características compositivas y de contenido que la imagen original (véase **Figura 1**).

Figura 1

Comparación visual entre el estímulo fotográfico auténtico y su réplica sintética generada mediante Magnific AI



Nota. En el panel izquierdo se presenta la fotografía auténtica de referencia. En el panel derecho se exhibe su contraparte sintética, cuya creación se basó en un prompt descriptivo extraído directamente de la original.

Para evaluar los patrones de percepción humana, el instrumento integró una sección de indicios visuales. Tras clasificar un estímulo como generado por IA, el formulario desplegaba una pregunta de opción múltiple que obligaba al participante a justificar su decisión eligiendo entre ocho variables de análisis: Detalle, Color, Suavidad, Sombras y luces, Desenfoque, Manchas o texturas mal aplicadas, Racionalidad o coherencia, e Intuición.

Procedimiento de recolección de datos

La recolección se estructuró de forma rigurosa para asegurar el control de las condiciones de aplicación del estudio:

Planificación y pilotaje: Diseño del banco visual y validación de la carga cognitiva en la plataforma virtual Google Forms para asegurar el correcto funcionamiento del instrumento.

Abordaje y autorización docente: El acceso a las aulas se coordinó de manera directa. El investigador se dirigió presencialmente a los cursos en las distintas facultades y solicitó la autorización verbal del docente en turno, explicando brevemente la finalidad del estudio para que permitiera la participación de los alumnos durante su espacio de clases.

Aplicación presencial supervisada: Una vez permitido el ingreso al aula, el levantamiento de datos se hizo de manera inmediata. A pesar de utilizar una herramienta en línea como Google Forms, el investigador permaneció de forma presencial para supervisar que los estudiantes respondieran desde sus dispositivos ya fuesen teléfonos o computadoras de manera individual, autónoma y sin recurrir a herramientas externas de verificación. La prueba se administró sin límite de tiempo, promoviendo una observación minuciosa de las imágenes.

Limitaciones del estudio

Entorno de aplicación: El despliegue masivo en aulas expuso a los participantes a distracciones y ruidos propios del contexto educativo compartido.

Variabilidad técnica de los dispositivos: Al responderse desde pantallas personales, resultó imposible estandarizar la resolución, el brillo y la calibración de color de cada dispositivo, factores tecnológicos que inciden en el descubrimiento de los detalles fotorrealistas.

Sesgo de deseabilidad social: El entorno presencial supervisado pudo inducir en ciertos participantes el deseo de proyectar mayor agudeza o competencia tecnológica.

Análisis de datos

Para evaluar objetivamente la habilidad de discriminación visual y aislar el sesgo cognitivo del participante, se aplicó la Teoría de Detección de Señales (TDS). Este marco analítico, diseñado para medir decisiones bajo condiciones de incertidumbre (Macmillan & Creelman, 2005), operó mediante una matriz 2x2 donde la señal a detectar era la imagen generada por IA y el "ruido" era la fotografía real. De las 50 respuestas de cada individuo, se extrajeron los conteos de Aciertos (H = era IA y fue clasificada como "IA"), y las Falsas Alarmas (FA = era Real y fue clasificada como "IA").

Para evitar que los valores extremos (0% o 100% de aciertos o falsas alarmas) indefinan el cálculo de las funciones inversas, se aplicó la corrección log-linear de Hautus (1995). Esta estandarización se ejecutó de forma uniforme en la base de datos de todos los participantes, sumando 0.5 al numerador y 1 al denominador de las tasas base:

$$\text{Hit Rate}_{\text{ajustada}} = \frac{H + 0.5}{25 + 1}$$

$$\text{FA Rate}_{\text{ajustada}} = \frac{FA + 0.5}{25 + 1}$$

A partir de estas probabilidades, se calcularon dos parámetros psicofísicos. El índice de detectabilidad (d'), que evalúa la sensibilidad real del participante, se obtuvo calculando la diferencia entre las transformaciones probit de ambas tasas:

$$d' = z(\text{Hit Rate}) - z(\text{FA Rate})$$

Donde z corresponde a la función inversa de la distribución normal acumulada estándar.

Para la interpretación de los resultados del índice d' , se empleó la escala orientativa basada en las convenciones clásicas de la psicofísica y la psicología experimental (Green y Swets, 1966; Macmillan y Creelman, 2005) (véase **Tabla 2**)

Tabla 2

Escala orientativa de d'

d'	Interpretación
≈ 0	Sin discriminación — la persona adivina, como tirar una moneda
0.2 – 0.5	Discriminación muy débil
0.5 – 1.0	Discriminación débil-moderada
1.0 – 2.0	Discriminación moderada
2.0 – 3.0	Discriminación buena
> 3.0	Discriminación excelente (casi sin superposición entre las distribuciones de "real" e "IA")

Nota. Escala orientativa de uso común para los valores del índice de detectabilidad (d'). Basado en los textos de la Teoría de Detección de Señales de Green y Swets (1966); Macmillan y Creelman (2005). Distintos autores mueven un poco los límites, así que se deben tratar como zonas aproximadas, y no como fronteras exactas.

Adicionalmente, se calculó el criterio o sesgo de respuesta (c) mediante la fórmula:

$$c = -0.5[z(\text{Hit Rate}) + z(\text{FA Rate})]$$

Un valor $c = 0$ indica ausencia de sesgo; valores $c > 0$ revelan un criterio conservador (necesidad de estar muy convencido para declarar "IA"); y valores $c < 0$ señalan un criterio liberal (tendencia a reportar "IA" ante la mínima duda).

Por último, para determinar el porcentaje general de aciertos de cada individuo, se calculó la Exactitud global (ACC) mediante la proporción de decisiones correctas sobre el total de estímulos:

$$ACC = \frac{H + CR}{50}$$

Donde H corresponde a los Aciertos y CR a los Rechazos Correctos, dividiendo la suma entre los 50 estímulos visuales evaluados por cada participante.

El análisis de los indicios visuales se procesó directamente sobre la base de datos de las respuestas mediante Microsoft Excel. Dado el formato de selección múltiple, las cadenas de texto se desglosaron para contabilizar las frecuencias absolutas a lo largo de las respuestas afirmativas a la detección de IA. En esta fase, la variable de Exactitud global (ACC) adquirió un rol metodológico clave al utilizarse como filtro de segmentación analítica. Tras contabilizar las frecuencias absolutas de las categorías visuales seleccionadas, se calcularon sus proporciones relativas dividiendo la data en dos escenarios de contrastación: el comportamiento de la población general sobre un total de razones y el comportamiento exclusivo de un subgrupo de alto rendimiento, construido filtrando estrictamente a aquellos participantes que alcanzaron un ACC superior al 70%.

Finalmente, las variables tabuladas se exportaron al software IBM SPSS Statistics para su respectiva contrastación. Como paso fundamental para la selección de las pruebas estadísticas, se determinó aplicar sistemáticamente las pruebas de

normalidad de Kolmogorov-Smirnov y Shapiro-Wilk sobre la totalidad de los resultados obtenidos, abarcando tanto los índices psicofísicos globales (d' , c , ACC) como las distribuciones de frecuencias y las métricas asociadas a los indicios visuales. La regla de decisión estadística estableció que, si los valores de significancia resultaban inferiores al umbral ($p < .05$), se rechazaría el supuesto de normalidad. Esta directriz procedimental configuró el criterio principal para justificar el uso de estadísticos inferenciales no paramétricos. Se establece como limitación explícita del estudio la posible dependencia intraclase derivada del muestreo por conglomerados de aulas, lo cual acota el alcance de las inferencias. En consecuencia, la ruta analítica planificó inicialmente una prueba t de Student para una muestra (asumiendo su robustez paramétrica ante el tamaño muestral como supuesto procedimental) para contrastar la sensibilidad perceptiva contra el valor nulo de azar absoluto ($d' = 0$); seguida de la prueba U de Mann-Whitney para evaluar las diferencias según el género y la experiencia previa con IA, y la prueba de Kruskal-Wallis para contrastar las variaciones de la capacidad perceptiva entre las distintas facultades y semestres académicos de la institución.

Resultados

Capacidad general para distinguir imágenes generadas por IA

Para dar respuesta al objetivo principal del estudio, se evaluó la sensibilidad de los estudiantes ($N = 381$) para discriminar entre imágenes fotorrealistas auténticas y aquellas generadas algorítmicamente, utilizando la Teoría de Detección de Señales. Los estadísticos descriptivos mostraron una media de $d' = 0.63$ (desviación estándar $[DE] = 0.49$; intervalo de confianza $[IC]$ del 95% $[0.59, 0.68]$). En relación con las escalas interpretativas, este valor se considera como una discriminación débil-

moderada. También, el análisis del sesgo de respuesta arrojó un promedio de $c = 0.61$ ($DE = 0.48$; IC 95% [0.57, 0.67]), lo cual indica un criterio conservador; es decir, los participantes mostraron una fuerte predisposición a designar los estímulos como reales, necesitando estar muy convencidos para clasificarlos como generados por IA. La prueba t de Student para una muestra evidenció que la sensibilidad global es significativamente superior al azar absoluto ($t(380) = 25.10$, $p < .001$). Sin embargo, al no alcanzar una discriminación moderada ($d' = 1.0$), se confirma una vulnerabilidad visual generalizada frente a las herramientas generativas.

Diferencias en la capacidad de detección según experiencia previa y género

En relación con los objetivos específicos, se analizó la influencia de las variables sociodemográficas y experiencia previa en el nivel de detección. Respecto a la experiencia previa en la creación de imágenes mediante inteligencia artificial, la mayor parte de la muestra afirmó poseer experiencia ($n = 316$, mediana = 0.63, rango intercuartílico [RIC] = 0.73) frente a una minoría que indicó no tenerla ($n = 65$, mediana = 0.47, RIC = 0.67). La prueba no paramétrica U de Mann-Whitney no evidenció diferencias estadísticamente significativas en la capacidad de detección entre ambos grupos ($U = 8886$, $p = .087$, tamaño del efecto $r = -0.088$). Aunque la falta de significancia no demuestra una equivalencia estadística perfecta entre grupos, los datos sugieren que la familiaridad técnica o el uso previo de estas plataformas no otorga una ventaja perceptiva real frente al fotorrealismo algorítmico. Al analizar las diferencias según el género de los participantes, los resultados mostraron valores muy similares entre mujeres ($n = 232$, mediana = 0.64, RIC = 0.81) y hombres ($n = 149$, mediana = 0.59, RIC = 0.63). La prueba U de Mann-Whitney no detectó diferencias estadísticamente significativas ($U = 16653.5$, $p =$

.548, $r = -0.031$), sugiriendo que la vulnerabilidad ante el engaño visual afecta de manera transversal a la muestra, sin distinción de género (véase Tabla 3).

Tabla 3
Estadísticos descriptivos y prueba U de Mann-Whitney para el índice de detectabilidad según género y experiencia previa

Variable	n	Mediana (d')	RIC	U	P	R
Género				16653.5	.548	-.031
Femenino	232	0.64	0.81			
Masculino	149	0.59	0.63			
Experiencia previa con IA				8886	.087	-.088
Sí posee experiencia	316	0.63	0.73			
No posee experiencia	65	0.47	0.67			

Nota. n = tamaño de la muestra; d' = índice de detectabilidad; RIC = rango intercuartílico; U = estadístico de Mann-Whitney; p = nivel de significancia; r = tamaño del efecto (coeficiente de correlación r derivado de Z). El nivel de significancia se calculó mediante la prueba no paramétrica U de Mann-Whitney para muestras independientes. Un valor $p < .05$ indica diferencias estadísticamente significativas. Un resultado no significativo falla en rechazar la hipótesis nula, pero no demuestra una equivalencia técnica perfecta entre los grupos.

Análisis complementario por facultad y nivel académico

De forma adicional, se exploró el rendimiento de los estudiantes según su área de estudio y avance académico mediante la prueba de Kruskal-Wallis. Se hallaron diferencias estadísticamente significativas en el nivel de detección según la facultad ($H(6) = 25.54$, $p < .001$, tamaño del efecto $\eta^2 = 0.052$, calculado con base en el estadístico H). Las pruebas *post hoc* de Dunn con ajuste de Bonferroni demostraron que los estudiantes de la Facultad de Ciencias Sociales, Derecho y

Bienestar (mediana = 0.80) presentan un rendimiento estadísticamente superior a los de la Facultad de Educación y Turismo (p ajustado = .014) y la Facultad de Ciencias Administrativas, Contables y Comercio (p ajustado = .004); en esta facultad destacada, el 21.5% de sus integrantes superó el 70% de exactitud global (ACC).

Asimismo, se encontraron variaciones significativas según el semestre en curso ($H(9) = 18.82, p = .027, \eta^2 = 0.026$). Resulta notable el desempeño de los estudiantes de noveno semestre (mediana = 1.01), quienes superaron significativamente en los contrastes *post hoc* de Dunn a los de sexto (p ajustado = .011) y séptimo semestre (p ajustado = .008), siendo el único grupo en alcanzar una discriminación moderada, con un 31.6% de sus integrantes superando el 70% de exactitud global (ACC).

Análisis de indicios visuales en la detección de imágenes sintéticas

Para dar respuesta al objetivo específico orientado a identificar los patrones de escrutinio visual, se analizaron las justificaciones perceptivas empleadas por los participantes al clasificar un estímulo como generado por inteligencia artificial. La evaluación contempló ocho indicios visuales predeterminados: *Detalle, Color, Suavidad, Sombras y luces, Desenfoque, Manchas o texturas mal aplicadas, Racionalidad o coherencia, e Intuición.*

El análisis descriptivo de las respuestas afirmativas reveló un total de 12.148 selecciones emitidas por la población general. Cabe recalcar que, al derivar de preguntas de opción múltiple, estas selecciones no constituyen observaciones estadísticamente independientes. Al examinar este volumen de datos, se observó que la estrategia de detección del estudiante promedio se basó predominantemente en la identificación de anomalías relacionadas con el **Detalle** (representando el 19.81% de las selecciones totales), seguido por el criterio de **Color** (15.37%). Resulta llamativo

que la **Intuición** ocupó una posición relevante en la población general, abarcando el 11.74% de las decisiones, lo que sugiere una tendencia a la adivinación o la falta de rigor técnico al momento de justificar el engaño visual.

Sin embargo, para aislar los patrones de observación empíricamente efectivos, se contrastó el comportamiento de la población general frente a un subgrupo de alto rendimiento. Este segmento poblacional fue filtrado bajo el criterio de poseer una Exactitud Global superior al 70% ($ACC > 70\%$), un umbral analítico de carácter exploratorio establecido para delimitar y estudiar exclusivamente al segmento con mayor capacidad de detección (el 13.12% superior de la distribución muestral). Este subgrupo registró un total de 1.019 selecciones justificativas. El análisis mostró que los sujetos con mayor precisión visual centran su enfoque en variables técnicas y reducen la dependencia heurística. En este subgrupo, el escrutinio de Detalle (20.22%) y Color (18.55%) mantienen su dominio con una mayor presencia relativa. Como contraparte, la justificación basada meramente en la Intuición experimentó una caída sustancial (representando apenas el 8.54% de sus selecciones) (véase **Tabla 4**).

Tabla 4
Comparación de Indicios Visuales entre población general y subgrupo de alto rendimiento

Indicio Visual (Razón)	Población General (n = 12,148)		Subgrupo de Alto Rendimiento (n = 1,019)	
	Frecuencia	%	Frecuencia	%
Detalle	2407	19.81%	206	20.22%
Color	1867	15.37%	189	18.55%
Suavidad	1696	13.96%	175	17.17%
Sombras y luces	1612	13.27%	133	13.05%
Intuición	1426	11.74%	87	8.54%
Desenfoque	1316	10.83%	112	10.99%

Manchas o texturas mal aplicadas	1024	8.43%	82	8.05%
Racionalidad o coherencia	800	6.59%	35	3.43%
Total	12.148	100%	1.019	100%

Nota. El subgrupo de alto rendimiento corresponde a los participantes que alcanzaron una Exactitud Global (ACC) superior al 70%. Se observa la reducción en la dependencia heurística.

A modo de síntesis, los hallazgos evidencian una vulnerabilidad generalizada ante el fotorrealismo algorítmico, caracterizada por un marcado sesgo conservador y un nivel de discriminación de magnitud débil a moderada. El análisis inferencial demostró que la experiencia técnica previa con herramientas generativas no otorga ventajas significativas para la detección. En contraste, el nivel académico constituyó un factor diferenciador, siendo los estudiantes de noveno semestre el único subgrupo en alcanzar una discriminación moderada. Finalmente, la segmentación exploratoria de la población con mayor rendimiento ($ACC > 70\%$) determinó que el patrón de evaluación de este grupo se asocia con una menor dependencia de la intuición y un aumento relativo en el escrutinio de categorías visuales como el Detalle, el Color y la Suavidad; presentando, por el contrario, una disminución en la identificación de Manchas o texturas mal aplicadas y en el análisis de la Racionalidad o coherencia. Esta evidencia empírica constituye la base para la discusión teórica subsecuente.

Discusión

La vulnerabilidad perceptiva frente a la hiperrealidad algorítmica

El hallazgo central de esta investigación revela una profunda vulnerabilidad en la percepción visual de la muestra de estudiantes evaluados. El índice de detectabilidad general ($d' = 0.63$) y la predisposición a clasificar el contenido sintético como auténtico (criterio conservador $c = 0.61$) evidencian que, frente al

fotorrealismo de las plataformas generativas actuales, la capacidad de discriminación en los participantes es estadísticamente superior al azar ($p < .001$), aunque en términos prácticos su magnitud refleja un nivel de discriminación débil a moderado. Estos resultados presentan una convergencia con la evidencia empírica global; por ejemplo, bajo condiciones metodológicas similares centradas en plataformas generativas de vanguardia, el estudio original de Lu et al. (2023) reportó una exactitud humana promedio del 61.3%, mientras que Bray et al. (2023) registraron un nivel del 62.18%. Además, el fuerte sesgo conservador ($c = 0.61$) encontrado en nuestra muestra, donde los participantes asumen por defecto la autenticidad de la imagen, es consistente con lo advertido por Velásquez-Salamanca et al. (2025), quienes encontraron que casi el 40% de las imágenes generadas por IA son clasificadas erróneamente como reales en entornos de escrutinio cotidiano. Todo esto respalda la postura teórica de Sánchez Díaz (2024) respecto a que el fotorrealismo algorítmico explota las heurísticas de confianza del cerebro bajo la premisa evolutiva de que "ver es creer", vulnerando la capacidad del individuo para discernir la verdad objetiva.

El mito del "usuario experto": Experiencia vs. Alfabetización

Un resultado estadísticamente revelador del presente estudio fue la ausencia de diferencias significativas entre los estudiantes que afirmaron tener experiencia previa en la creación de imágenes mediante IA ($n = 316$) y aquellos sin experiencia ($n = 65$). Este fenómeno guarda concordancia con los hallazgos cuantitativos de Lu et al. (2023), quienes reportaron que el poseer experiencia previa con herramientas generativas otorgaba una ventaja perceptiva marginal (+3.7%) que no resultó estadísticamente determinante. En el presente estudio, si bien el análisis no paramétrico no arrojó diferencias significativas, este resultado debe interpretarse con

prudencia debido a la marcada asimetría de los grupos (316 frente a 65), lo cual no prueba una equivalencia total ni la ausencia absoluta de un efecto, sino que sugiere como hipótesis que el uso funcional o cotidiano de estas plataformas no garantiza el desarrollo de un blindaje perceptivo superior. Esto respalda la postura de García-Delgado Giménez y Revilla Guijarro (2025), quienes argumentan que la alfabetización mediática no debe ser funcional, sino conceptual. Saber usar un modelo generativo no implica necesariamente poseer las habilidades críticas requeridas para evaluar cómo estas estructuras alteran la percepción de la realidad.

El impacto de la formación académica y las asimetrías de la muestra

Al explorar el rendimiento por facultades y niveles académicos, los datos arrojaron variaciones significativas que requieren una interpretación prudente debido a la distribución asimétrica de los participantes. El bajo rendimiento registrado en áreas como la Facultad de Artes, Humanidades y Patrimonio ($d' = 0.29$) debe matizarse frente a su reducido tamaño muestral ($n = 8$). En contraste, se registró un desempeño destacado en la Facultad de Ciencias Sociales, Derecho y Bienestar ($d' = 0.81$, $n = 79$) y en la Facultad de Ciencias de la Salud ($d' = 0.69$, $n = 106$). El rendimiento de estas áreas, sumado al pico de detectabilidad observado en los estudiantes de noveno semestre ($d' = 1.01$), concuerda con la necesidad expuesta por García-Delgado Giménez y Revilla Guijarro (2025) de transformar la alfabetización mediática en una competencia transversal. No obstante, al tratarse de un diseño transversal, este resultado no permite atribuir el rendimiento exclusivamente a la formación académica, ya que el pensamiento crítico frente al engaño cognitivo también podría estar mediado por la edad, la exposición tecnológica previa o la composición misma de la muestra.

De la intuición al escrutinio metódico: estrategias de lectura visual

La exploración de los indicios visuales sugiere empíricamente que la detección de contenido sintético requiere de un proceso de evaluación técnica. A nivel descriptivo, el hecho de que la población general justificara un porcentaje relevante de sus decisiones (11.74%) basándose puramente en la "Intuición", podría reflejar una lectura holística y superficial. Esta tendencia concuerda con lo observado por Högemann et al. (2025), quienes documentaron que los usuarios que se guiaban por la intuición (juzgando las imágenes como "demasiado perfectas") tendían a disminuir su precisión frente a modelos generativos avanzados. Por el contrario, la exploración descriptiva del subgrupo de alto rendimiento ($ACC > 70\%$) mostró una reducción aparente en esta dependencia intuitiva (8.54%) para concentrar su atención en variables estructurales como el "Detalle" (20.22%) y el "Color" (18.55%). Este enfoque de lectura se asimila a los hallazgos de Lu et al. (2023), quienes identificaron precisamente las anomalías de "detalle" y "suavidad" como las principales vulnerabilidades algorítmicas detectables por el ojo humano. Este cambio para deconstruir la imagen se alinea con la diferenciación ontológica de Marín Álvarez et al. (2025), respaldando que las imágenes sintéticas son ensamblajes probabilísticos carentes de anclaje lumínico o material, cuyas inconsistencias solo pueden ser reveladas mediante un escrutinio estructural focalizado.

Futuras líneas de investigación

A partir de los hallazgos de esta investigación, resulta imperativo trazar nuevas rutas de indagación científica. En primer lugar, el pico de rendimiento del noveno semestre sugiere la necesidad de realizar estudios que determinen cómo evoluciona la capacidad de discriminación visual en los estudiantes universitarios

desde que comienzan hasta finalizar su formación académica. En segundo lugar, debido al rápido avance de la IA generativa, es necesaria una comparación entre segmentos por arquitecturas de generación específicas, como modelos generativos antagonísticos frente a modelos de difusión, con el fin de evaluar si el rigor técnico centrado en el "Detalle" y el "Color" seguirá siendo factible frente a la inminente pérdida de anomalías visuales en las nuevas versiones algorítmicas. Finalmente, sería valioso replicar la aplicación de este instrumento aislando la variable de especialización disciplinar, contrastando la percepción de la población universitaria general frente a profesionales activos del área de la comunicación y medios visuales, para corroborar si la experiencia profesional logra mitigar el sesgo de confianza y la vulnerabilidad cognitiva demostrada en este estudio.

Conclusiones

La presente investigación evidencia que existe una vulnerabilidad visual en la muestra analizada de 381 estudiantes frente al banco de estímulos de imágenes generadas por inteligencia artificial. La habilidad de los participantes, bajo las condiciones de aplicación del estudio, para diferenciar entre una fotografía real y una sintética evidenció una fuerte predisposición a confiar en la veracidad de la imagen antes que dudar de ella. Estos hallazgos sugieren que el fotorrealismo de las plataformas generativas actuales es capaz de vulnerar los límites perceptivos de la población evaluada.

En cuanto a los aspectos sociodemográficos y técnicos, se concluye que no se detectaron diferencias estadísticamente significativas en esta muestra según el género o la experiencia previa utilizando herramientas generativas. Estos resultados sugieren que saber operar funcionalmente una plataforma de inteligencia artificial no

se traduce necesariamente en poseer el escrutinio crítico requerido para auditar la veracidad de sus resultados visuales.

No obstante, los resultados sugieren una asociación positiva entre ciertas áreas de estudio y un mejor desempeño perceptivo. El rendimiento significativamente superior de los estudiantes de noveno semestre y de aquellos pertenecientes a la Facultad de Ciencias Sociales, Derecho y Bienestar indica que podría existir una relación con metodologías de escrutinio analítico; sin embargo, al tratarse de un diseño transversal, este hallazgo no puede establecer una relación de causa y efecto directa, ya que no se aisló el efecto exclusivo de la formación académica frente a otras variables de confusión.

Finalmente, a nivel exploratorio, se observa que en la muestra estudiada la detección de IA generativa no surgió de forma instintiva. Mientras que un porcentaje de los participantes basó sus justificaciones en la mera intuición, el subgrupo de alto rendimiento adoptó un patrón de escrutinio visual más técnico. Los datos sugieren que un análisis efectivo parece requerir el abandono de la lectura holística de la fotografía para deconstruir la imagen y examinar posibles inconsistencias algorítmicas, concentrándose específicamente en la aplicación del detalle y el color de los estímulos evaluados.

Referencias

- Activa Research, & Worldwide Independent Network of Market Research. (2023). *Fake news: Desinformación en Chile y LatAm. Estudio Desinformación*. https://chile.activasite.com/wp-content/uploads/2023/04/222287_Estudio-Desinformacio%CC%81n_0329_V1-2.pdf
- Altamirano Molina, M. V. (2026). Deepfake en Ecuador: Una mirada desde la cobertura periodística. *Uru. Revista de Comunicación y Cultura*, 13, 190-206. <https://doi.org/10.32719/26312514.2026.13.11>
- Bray, S. D., Johnson, S. D., & Kleinberg, B. (2023). Testing human ability to detect 'deepfake' images of human faces. *Journal of Cybersecurity*, 9(1), tyad011. <https://doi.org/10.1093/cybsec/tyad011>
- Casu, M., Guarnera, L., Zangara, I., Caponnetto, P., & Battiato, S. (2025). A (Mid)journey through reality: Assessing accuracy, impostor bias, and automation bias in human detection of AI-generated images. *Human Behavior and Emerging Technologies*, 2025, 9977058. <https://doi.org/10.1155/hbe2/9977058>
- Cruz-Meléndez, C. (2025). La manipulación de la evidencia de políticas públicas con la Inteligencia Artificial Generativa: los riesgos de los deepfakes. *Universitas XXI*, 43, 15-36. <https://doi.org/10.17163/uni.n43.2025.01>
- Diel, A., Lalgı, T., Schröter, I. C., MacDorman, K. F., Teufel, M., & Bäuerle, A. (2024). Human performance in detecting deepfakes: A systematic review and meta-analysis of 56 papers. *Computers in Human Behavior Reports*, 16, 100538. <https://doi.org/10.1016/j.chbr.2024.100538>
- El Salous Muñoz, A., Naranjo Muñoz, P., Collaguazo Mejía, M., & Rodríguez Rivas, J. (2026). Inteligencia artificial y desinformación: Un análisis de la

percepción y credibilidad del contenido sintético en una universidad técnica del Ecuador. Escuela Superior Politécnica del Litoral.

https://www.researchgate.net/publication/399875153_Inteligencia_artificial_y_desinformacion_Un_analisis_de_la_percepcion_y_credibilidad_del_contenido_sintetico_en_una_universidad_tecnica_del_Ecuador

- García-Delgado Giménez, B., & Revilla Guijarro, A. (2025). Alfabetización mediática frente a la desinformación: revisión de propuestas educativas en el ámbito universitario hispanohablante (2014–2023). *Dígitos. Revista de Comunicación Digital*, 11, 78-99. <https://doi.org/10.7203/drdcd.v0i11.313>
- García-Marín, D., Suárez-Álvarez, R., & García-Jiménez, A. (2025). "Todo parece veraz". Credibilidad de la desinformación producida usando IA desde la perspectiva de los estudiantes de comunicación en España. *Revista de Comunicación*, 24(2), 183-227. <https://doi.org/10.26441/RC24.2-2025-3872>
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley.
- Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d'. *Behavior Research Methods, Instruments, & Computers*, 27(1), 46-51. <https://doi.org/10.3758/BF03203619>
- Högemann, M., Betke, J., & Thomas, O. (2025). What you see is not what you get anymore: a mixed-methods approach on human perception of AI-generated images. *Frontiers in Artificial Intelligence*, 8, 1707336. <https://doi.org/10.3389/frai.2025.1707336>
- Lu, Z., Huang, D., Bai, L., Qu, J., Wu, C., Liu, X., & Ouyang, W. (2023). Seeing is not always believing: Benchmarking human and model perception of AI-

- generated images. *Advances in Neural Information Processing Systems*, 36, 25435-25447. <https://doi.org/10.52202/075280-1105>
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2.^a ed.). Lawrence Erlbaum Associates.
- Marin Alvarez, M. A., Acero Gutiérrez, A., García-Hernández, M., & Hurtado Espinosa, C. L. (2025). Las imágenes creadas con Inteligencia Artificial no son fotografías. *Revista Caribeña de Ciências Sociais*, 14(10), 1-18. <https://doi.org/10.55905/rcssv14n10-016>
- Nightingale, S. J., & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. *Proceedings of the National Academy of Sciences*, 119(8), e2120481119. <https://doi.org/10.1073/pnas.2120481119>
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388-402. <https://doi.org/10.1016/j.tics.2021.02.007>
- Saeidnia, H. R., Jahani, S., Ghiasi, N., & Keshavarz, H. (2026). Generative AI and health misinformation: production, propagation, and mitigation—a systematic review. *BMC Public Health*, 26, 693. <https://doi.org/10.1186/s12889-025-26148-9>
- Sánchez Díaz, M. F. (2024). El impacto de la inteligencia artificial generativa en los derechos humanos. *Revista Eurolatinoamericana de Derecho Administrativo*, 11(1), e252. <https://doi.org/10.14409/redoeda.v11i1.13612>
- Universidad Laica Eloy Alfaro de Manabí. (2025). *Informe Preliminar de Rendición de Cuentas 2024*. <https://www.uleam.edu.ec/wp-content/uploads/2025/07/Informe-Preliminar-de-Rendicion-de-Cuentas-2024-ULEAM.pdf>

Vega Cocha, D. P., Vásquez Chicaiza, F. P., Guevara Guevara, A. M., Maldonado Palacios, I. A., & Maldonado Arce, M. E. (2024). Desinformación en la era digital: el papel de las redes sociales en la propagación de noticias falsas durante conflictos globales. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, 5(2), 138-149.

<https://doi.org/10.56712/latam.v5i2.1865>

Velásquez-Salamanca, D., Martín-Pascual, M. Á., & Andreu-Sánchez, C. (2025). Interpretation of AI-generated vs. human-made images. *Journal of Imaging*, 11(7), 227. <https://doi.org/10.3390/jimaging11070227>

World Economic Forum. (2025). *Top 10 emerging technologies of 2025*.

<https://www.weforum.org/publications/top-10-emerging-technologies-of-2025/>