



IMAGINARIO CULTURAL DE LAS FIESTAS DE CASTILLA Y LEÓN GENERADO CON INTELIGENCIA ARTIFICIAL Estudio comparado mediante Playground AI

MATIAS LÓPEZ IGLESIAS¹, JOSÉ L. CARREÑO VILLADA¹, SANDRA GONZÁLEZ PARRA¹

¹ Universidad Europea Miguel de Cervantes, España

PALABRAS CLAVE

*Patrimonio cultural
Visual
Fiestas patronales
Inteligencia artificial
Imagen por ordenador
Fotografía
Denotación
Connotación*

RESUMEN

Este estudio analiza la generación de 133 imágenes mediante el uso de inteligencia artificial (IA) para representar Fiestas de Interés Cultural (FIC) de Castilla y León utilizando la plataforma Playground AI. Las imágenes se generaron a partir de fotografías reales de las fiestas. Los resultados se evaluaron mediante encuestas comparativas, centradas en dos aspectos: la familiaridad y confort que generan, y su parecido con la realidad. Aunque las fotografías originales tienen mayor semejanza con la realidad, las imágenes creadas por Playground AI destacan por representar mejor las fiestas y resultar más atractivas para los encuestados. Concluimos que estas herramientas digitales adaptan tanto la dimensión simbólica como la literal, conservando la esencia de la fuente original. Además, uniforman la estética mediante filtros artísticos con armonía cromática, logrando una apariencia visual más agradable y adecuada para destacar los aspectos representativos de estas celebraciones culturales.

Recibido: 27/ 05 / 2024
Aceptado: 17/ 09 / 2024

1. Introducción

La inteligencia artificial (IA) es el nombre que se le asigna a una serie de tecnologías con características o capacidades que antes eran exclusivas del intelecto humano. El término se aplica cuando una máquina imita las funciones cognitivas que los humanos asocian con otras mentes humanas, como aprender o resolver problemas (Icarte Ahuada, 2016; Toca Torres, 2014).

La inteligencia artificial está cada vez más presente en la vida diaria de las personas y es que se interactúa con ellas para realizar actividades cotidianas como, por ejemplo, escuchar música, apagar las luces monitorizadas de casa o encontrar una ruta para llegar a un destino (Cwaik, 2020). Hemos recurrido a la inteligencia artificial para crear una serie de imágenes que ilustran y acompañan la descripción de las Fiestas de Interés Regional, Nacional e Internacional de Castilla y León. La necesidad de abordar una representación visual de estas fiestas confronta con los problemas legales de la autoría de las fotografías. Los problemas legales sobre la autoría de la inteligencia artificial están desafiando las normas tradicionales de los derechos de propiedad, ya que los algoritmos pueden generar obras que podrían considerarse creativas si fueran producidas por humanos (Estupiñán Ricardo et al., 2021).

1.1. El valle inquietante en la cultura visual: explorando la algoritmografía

Masahiro Mori, un japonés experto en robótica, ideó la teoría del «valle inquietante» en 1970. Esta teoría explica que los robots son agradables cuando se parecen a los humanos, pero al alcanzar una gran semejanza, despiertan un miedo natural. Mori usó la metáfora del valle para describir cómo aumentar los rasgos humanos en algo no humano genera una sensación de extrañeza y alejamiento de nuestra simpatía. Esta teoría se extendió a cualquier figura antropomórfica, explorando nuestra incomodidad y rechazo hacia representaciones físicas y virtuales similares a los humanos, cuando no llegan a alcanzar una semejanza cien por ciento fiel a la realidad (Milo, 2023) como por ejemplo un robot. Estudios han demostrado que los robots resultan inquietantes cuando se perciben como capaces de tener emociones humanas, ya sea por su apariencia, sus acciones o simplemente por cómo se describen. Una explicación para este fenómeno es que los robots humanoides pueden resultar perturbadores porque parecen poseer una mente similar a la humana. Aunque los robots no tienen una mente real, muchas personas perciben características mentales en ellos, lo que puede generar sentimientos de malestar. Esta percepción de la mente en los robots humanoides se debe a la atribución de características humanas, como la capacidad de sentir emociones, lo que puede generar aversión afectiva. Esta percepción mental en los robots humanoides puede llevar a una discrepancia entre las expectativas de las personas y el conocimiento explícito de que los robots son simplemente máquinas, lo que genera malestar emocional.

En el contexto del marketing digital, la popularidad creciente de herramientas de inteligencia artificial, como ChatGPT, plantea desafíos similares. Aunque estas herramientas buscan imitar el lenguaje humano para crear contenido en línea, deben equilibrar la efectividad con la evitación del valle inquietante. Los usuarios pueden sentirse incómodos si perciben que están interactuando con contenido generado por IA que parece demasiado humano, lo que puede afectar su percepción del mensaje y la marca detrás de él (Lizama, 2024).

Las imágenes ilustrativas de cultura visual que se han generado mediante IA pueden presentar, también, este problema de inquietud en el espectador. Se ha buscado esquivar dicho valle inquietante para no generar malestar o incomodidad en los espectadores. Para ello se ha recurrido a un *prompt*, instrucciones de generación, solicitando a las herramientas un estilo pictórico a dichas imágenes, que alejan a los personajes, figuras o paisajes aún más de la realidad.

Es complicado generar imágenes con un contexto cultural. La inteligencia artificial (IA) está transformando el proceso de creación y significación. Hernando Gómez Gómez y José Luis Rubio Tamayo (2023) introducen el concepto de *algoritmografía* como proceso de creación de imágenes a partir de otras imágenes. Se destaca la importancia de considerar la dimensión ecológica de la imagen en el contexto de su producción y consumo en la sociedad actual. Además, exploran nociones de adopción y apropiación en la creación e interpretación de imágenes. Se argumenta que la adopción y apropiación adquieren una nueva dimensión con la introducción de las IA, que pueden generar imágenes basadas en conceptos y significados previamente definidos por la colectividad.

La situación actual está marcando la deriva de la comunicación tanto en la generación textual artificial como de imágenes. Un hecho que resulta inquietante por la forma en la que los especialistas en comunicación se deben enfrentar a la creación de material y a su capacidad para distinguir lo real de

aquello que ha sido creado bajo el prisma de una IA y su percepción pública, como aseguran Gómez de Ágreda et al. (2021) este tipo de percepciones «están cada vez más configuradas desde medios digitales susceptibles de ser manipulados por procesos sofisticados de inteligencia artificial (Lin, 2018) que permiten modificar imágenes de maneras antes prácticamente imposibles» (p. 3).

1.2. La semántica visual: de Peirce y Barthes a las inteligencias artificiales

Para generar las imágenes debemos partir de la semántica y semiótica (Mengoni, 2021), destacando la evolución histórica de la comprensión y representación visual. Se hace referencia a la teoría de Peirce sobre los signos (Bravo Heredia, 2010), ahondando en cómo influye en la forma en que las inteligencias artificiales y los algoritmos analizan y construyen nuevas imágenes, incorporando elementos culturales. Gómez Gómez y Rubio Tamayo (2023) argumentan que las categorías semióticas de Peirce son fundamentales para comprender la evolución de los significados de las imágenes y la construcción narrativa que las rodea. La relación entre las IAs y la teoría de categorías de Peirce sugiere que las IAs están aprendiendo gradualmente a entender y generar conceptos asociados a las imágenes.

Se discute la naturaleza de las imágenes generadas por las IAs, destacando su aproximación abstracta a la representación de fiestas de interés turístico y la importancia de considerar su contexto cultural y folclórico. Es necesario un análisis de la significación, denotativa y connotativa, de la imagen. con el fin de comprender mejor la relación entre semántica y representación visual, dado que las IAs y otras tecnologías pueden contribuir a generar visualmente imágenes artificiales.

No es necesario que haya una equivalencia figurativa directa entre la imagen y su referencia original. Justo Villafañe y Norberto Mínguez (2002) proponen un modelo de grados de iconicidad equivalencia texto-imagen-concepto, que incluye la equivalencia literal, simbólica, cromática entre otros. Diversos autores como Gómez Gómez y Rubio Tamayo (2023), Jordanous (2012), López Cardona et al. (2020) o McCormack et al. (2020), analizan el impacto de la inteligencia artificial, así como la posibilidad de que las IAs desarrollen habilidades artísticas similares a las humanas. Es una realidad que las IAs continuarán explorando nuevas formas de generar imágenes, pero la fotografía seguirá siendo importante como medio de capturar la realidad visual.

Roland Barthes (Menéndez-Pidal, 2010) identifica la oposición entre denotación y connotación en tres órdenes: codificado frente a no codificado, perceptivo frente a cultural y literal frente a simbólico. De los dos mensajes icónicos, el denotativo sirve de base al segundo: el mensaje literal actúa como fundamento del mensaje no codificado. La imagen literal representa lo denotado, mientras que la simbólica, lo connotado. El espectador percibe tanto el mensaje visual como el cultural simultáneamente.

Si atendemos a la simplificación de conceptos Barthes (Allen, 2011) entendía que la intertextualidad comprende la relación entre textos, es decir que los textos están entrelazados en una red de referencias y significados. Una IA con capacidad de aprendizaje usará los textos e imágenes que le han sido facilitados para generar un nuevo discurso o una nueva interpretación de la realidad; su red neuronal artificial interrelaciona los textos proporcionados y generará uno nuevo con una relación directa a los facilitados.

Como se puede observar los activos académicos pasados son el cimiento de una nueva reconfiguración de la realidad donde la autoría de los mensajes, en algunos casos, es difícil de distinguir, pero se trata de un contexto donde el proceso de adaptación debe ser gradual y atender a límites éticos y, por supuesto, legislativos.

1.3. Términos Clave en el análisis de imágenes generadas por inteligencia artificial

Hernando Gómez Gómez y José Luis Rubio Tamayo (2023) proponen un modelo de niveles de equivalencia entre texto, imagen y concepto, en el cual las inteligencias artificiales pueden generar imágenes con distintos niveles de interpretación y abstracción del lenguaje. Estos tres primeros niveles son:

- Equivalencia figurativa literal: Se refiere a la similitud de acciones y disposición de elementos en la imagen con respecto al concepto original, lo que conlleva una historia similar.

- Equivalencia figurativa simbólica: Aquí, algunos elementos pueden variar, pero la significancia del relato de la imagen permanece similar, requiriendo comprensión del relato original.
- Gama cromática: La modificación del color puede alterar ciertos niveles del mensaje y la dimensión simbólica de la imagen, dependiendo de factores culturales.

A lo que debemos sumar en el análisis la denotación y la connotación como significado de la imagen (Menéndez-Pidal, 2010):

- Significado denotativo: objetivo, literal. La palabra significa lo mismo que marca el diccionario. La denotación no deja lugar a segundas interpretaciones. Expresa una realidad exacta del término.
- Significado connotativo: subjetivo, simbólico. Dependiente del contexto, transmite sentimientos y emociones. Puede verse fuertemente condicionado por las costumbres o experiencias de la región y cultura.

1.4. Objetivos y preguntas de investigación.

El objetivo general de este estudio es comprobar si la teoría del Valle Inquietante de Masahiro Mori aplica también a las representaciones culturales creadas por inteligencia artificial. Específicamente, se busca evaluar en qué medida estas imágenes generadas logran sortear el Valle Inquietante mediante la inclusión de recursos técnicos y estéticos (OE1). Para ello, es indispensable verificar si la retórica visual planteada por Barthes se integra en los procesos generativos de la IA (OE2).

A la hora de plantear esta investigación, surgen una serie de preguntas que deben ser resueltas. (P1) ¿Cómo se representa visualmente la composición de personajes y escenarios de las fiestas culturales, lo que hace referencia al nivel de equivalencia figurativa literal? (P2) ¿Cómo se representa la fiesta y su construcción mental en el nivel de equivalencia figurativa simbólica? Además, (P3) ¿armonizan las imágenes generadas por IA la luz y el color, abordando la gama cromática de las imágenes?

También, nos preguntamos (P4) ¿cuál transmite mejor sensación de afinidad o confort? y (P5) ¿cuál de las representaciones generadas por IA se parece más a la realidad? Este análisis permitirá comprender mejor cómo la inteligencia artificial puede replicar y mejorar las representaciones culturales, manteniendo una estética agradable y evitando las sensaciones incómodas asociadas con el Valle Inquietante.

En cuanto a los significados denotativos y connotativos, (P6) ¿las recreaciones digitales representan la realidad de manera objetiva, literal e informativa?, o (P7) ¿prevalece más el significado connotativo, analizando los componentes subjetivo, simbólico y emocional? Ambos aspectos están intrínsecamente relacionados con el nivel de parecido humano al que hace referencia el Valle Inquietante.

2. Metodología

La muestra de análisis se ha generado mediante la aplicación de Generación de imágenes Playground AI (versión 2023). Se han creado un total de 133 ilustraciones artificiales finales que representen cada una de las 133 Fiestas de Interés Cultural registradas en distintos lugares de la región Castellano leonesa. Para comprobar su adecuación, se pidió la opinión a 266 encuestados acerca de la adecuada representación visual de cada festividad.

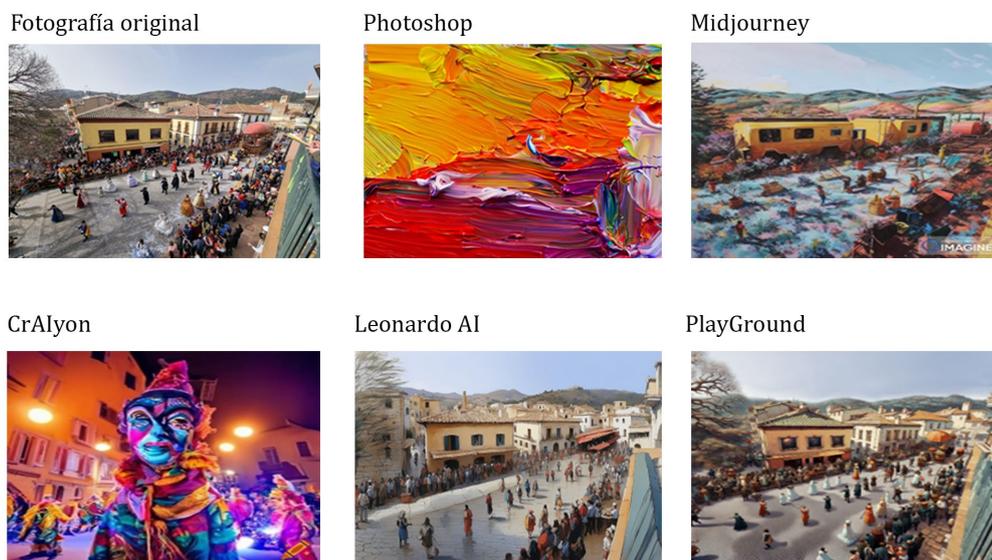
2.1. Generación de imágenes por IA

En primer lugar, se experimentó con distintos programas (Figura, 1) de inteligencia artificial que utilizan algoritmos para generar imágenes: Dall-e 2, Photoshop, Midjourney, Craiyon, Leonardo AI y Playground AI, decantándonos por este último. Se probó a generar imágenes con texto introduciendo una descripción de la fiesta, sin embargo, los resultados fueron tan poco aceptables que se decidió entrenar la aplicación con varias imágenes de las fiestas que se querían recrear. Hemos optado por utilizar esta herramienta para alejarnos de la realidad fotográfica y acercarnos a imágenes con texturas y formas pictóricas. Esto se debe, por una parte, a que no conseguimos obtener de todas las fiestas fotografías libres de derechos. En todas ellas, utilizamos una fotografía real de referencia.

La siguiente imagen de ejemplo hace referencia a Los Carnavales de Cebreros de Ávila (Ficha 37). El Ayuntamiento facilitó una serie de fotografías de la fiesta libres de derechos. En todos los casos se le

indicó el mismo *prompt*, instrucción de texto: «transforma en pintura al óleo, respeta la imagen, respeta sus elementos». Dependiendo de cómo analizase la IA la imagen de referencia, se añadieron descripciones más específicas en inglés porque observamos que el programa interpreta de una forma más efectiva las instrucciones suministradas en este idioma.

Figura 1. Imágenes generadas por Inteligencia Artificial de la Fiesta: Los Carnavales de Cebreros, Ávila.



Fuente: Elaboración propia, a partir de (de izquierda a derecha) fotografía original, Photoshop, Midjourney, CraiYon, Leonardo AI y Playground AI, 2024.

Se barajó usar la herramienta DALL-E 2, pero no nos permitió utilizar una imagen de base. Como se observa en la figura 1, Photoshop, aunque útil para retocar imágenes, no logró el efecto deseado al transformar la imagen en general.

Midjourney permite tomar una imagen de referencia y aplicar un estilo artístico, pero el resultado no cumplió con las expectativas, ya que la imagen se asemeja en cuanto a encuadre y formas, pero no respeta los elementos originales.

Craiyon, al igual que DALL-E 2, no permite el uso de una imagen de referencia y solo acepta descripciones de *prompt*. A pesar de intentarlo con instrucciones detalladas, el resultado no fue adecuado, ya que generaba varias imágenes relacionadas con la palabra principal «carnaval», algunas de las cuales eran demasiado alejadas y poco representativas de la fiesta.

Leonardo AI, nos permite establecer un *prompt* y subir una imagen de referencia, y ofreció resultados más prometedores, conservando colores, encuadre y estructura. Pero no mantenía los trajes de carnaval, que eran elementos importantes de la imagen original.

Finalmente, se optó definitivamente por operar con Playground AI, ya que, mediante el uso de *prompt* y la opción de entrenamiento mediante carga de ficheros «*On a Board*» utilizando en todos los casos el filtro llamado «*Lush Illumination*», se logró la transformación pictórica relativamente fiel a la imagen original de las fiestas analizadas. Aunque en cada ficha, fue necesario detallar más aspectos en el *prompt* para refinar el resultado, el proceso consistió en entrenar con la opción a partir de imágenes para lograr resultados homogéneos en cuanto a estilo, pero heterogéneos y adaptados al espíritu de cada festividad.

En algunas pruebas se modificó el estilo del filtro y no se encontraron cambios significativos en las imágenes, en el estilo, que es lo que dice cambiar. No se incluyó la opción «*Expand Prompt*» no la activamos. Este botón, al activarlo, hace que la IA improvise sobre el *prompt* que le has escrito, generando y sugiriendo imágenes e ideas. Si deseas ceñirte a una imagen de base, lo mejor es no activarlo, por lo que pudimos comprobar, ya que si no se aleja de la realidad e inventa elementos que guardan una relación objetiva de semejanza con aquello se le está pidiendo.

El proceso de entrenamiento más efectivo fue el de «*Image to Image*». El entrenamiento por *prompt* de texto en castellano, ofreciendo un párrafo resumen de cada fiesta no obtuvo los resultados esperados. En todos los casos se sustentó la IA mediante una serie de fotografías reales de las fiestas de cada lugar.

Se optó por el modelo «*Stable Diffusion XL*», con el fin de ofrecer más variedad en cuanto a posibilidades y dimensiones. El entrenamiento de IA se basa en un modelo matemático algorítmico que clasifica y reinterpreta los patrones (Heras Benavides, 2017) a través de la interrelación entre unidades, neuronas artificiales, conectadas (Tostado Sánchez *et al.*, 2016) que requiere del suministro de una gran cantidad de información a fin de poder obtener resultados óptimos para esta investigación.

En términos de tamaño y resolución es necesario tener en consideración que, a primera vista, lo que mejor puede parecer es la opción 1024x1024, ya que es la que mayores dimensiones posee y por lo tanto más calidad de imagen (Yoguel *et al.*, 2021). Sin embargo, después de diferentes pruebas se pudieron determinar que esta opción puede no resultar la más válida en algunos casos. Esto se debe a que la IA trata de estirar y hacer uso del relleno generativo todo ese espacio, deformando los elementos de la imagen. Por un lado, estira demasiado las figuras y el entorno y, por otro, incluye elementos poco realistas aun tratándose de una creación pictórica. Existe la posibilidad de editar dicha imagen para hacerla menos alargada, pero los resultados obtenidos dejaban entrever una imagen poco natural. Una propuesta más adecuada es elegir la dimensión 768x512 si lo que se busca es la creación de una imagen en horizontal. Con esta alternativa se consigue que una imagen no se vea deformada y rellena bien todos los espacios. Por el contrario, si se busca una imagen en vertical, otra alternativa factible es la elección del tamaño de 512x768.

Dependiendo de la imagen de referencia, se generaban mejor unas u otras: Las personas y algunos animales, toros o caballos, así como imágenes religiosas y procesionales, tallas de Vírgenes o Santos, no los reconoce adecuadamente. Lo mejor que generó son paisajes y edificios. Un ejemplo de ello es la festividad de la Romería de Nuestra Señora de Chilla, Candeleda, Ávila. Por ejemplo, La imagen de la Virgen (Figura 2) con el bebé en brazos, Playground AI no fue capaz de procesarla adecuadamente, además de los diferentes rostros de la gente a su alrededor, en cuanto a la forma y rasgos humanos, incluso dándole las instrucciones en inglés. Se recurrió, en la medida de lo posible a entrenar la IA con fotografía donde las personas que aparecieran en la foto estuvieran, en la medida de lo posible de espaldas a la cámara, para evitar infringir derechos de imagen y lograr mejores resultados.

Figura 2. Versiones de la Romería de Nuestra Señora de Chilla, Candeleda (Ávila, España)



Fuente: Elaboración propia, a partir Fotografías del Ayuntamiento de Ávila, 2024.

2.2. Realización de la encuesta

Para cotejar los resultados se han suministrado 266 fichas de fiestas de Interés Turístico Cultural (FIC) de Castilla y León a sujetos adultos abordados en la calle que voluntariamente accedían a rellenar la encuesta. Cada encuestado relleno una de las 133 fichas de fiestas de Interés Turístico de Castilla y León. Cada ficha consistía en un breve texto con las instrucciones impresas. El encuestador, presente en todo momento, aclaraba conceptos complejos como que se entiende por denotativo y connotativo, o cualquier duda que pudiera surgir.

Para dar más firmeza estadística se suministró cada ficha a dos personas. Cada ficha consistía en un breve texto con las instrucciones y dos imágenes, a la izquierda una fotografía real de la fiesta y a la derecha la ilustración generada por Playground AI. Exceptuando 3 encuestados que tardaron en realizarla más de 20 minutos, el tiempo medio fue de 5 minutos y 42 segundos.

En todas las fichas aparece en la parte superior las instrucciones con la URL y el código QR de la encuesta. A modo de ejemplo (Figura 3) se presenta la ficha número 52, que hace referencia a La Pingada de Mayo. Como se observa a la izquierda se encuentra una fotografía real tomada por Raúl de la Montaña

que se encuentra alojada en la web del Ayuntamiento de Hontoria del Pinar. A la derecha la imagen generada por la IA.

Figura 3. Ficha nº 52 / Fiesta: Fiesta del Mayo Hontoria del Pinar. (Soria, España).

Estudio sobre la **Fiestas de Interés cultural**. Por favor, siga las siguientes instrucciones:

1. Entre en el link de la encuesta <https://forms.office.com/e/YFzu9hMMfW> o escanee el código QR
2. Introduzca el nº de la imagen que aparece junto al título.
3. Lea detenidamente el título, y observe las dos imágenes que lo acompañan (Si no le queda claro la temática puede leer el resumen que lo acompaña).
4. Responda según su criterio a las siguientes 7 preguntas indicando cuál de las dos imágenes (fotografía e ilustración) se ajusta mejor a cada criterio.



52- Fiesta de Mayo, Hontoria del Pinar.



[Fotografía]



[Ilustración]

Fuente: Elaboración propia, a partir Foto original (de la Montaña, 2001), 2024.

Las preguntas realizadas comienzan consultando su grado de conocimiento, en formato de escala Likert, que tiene el encuestado de la fiesta analizada. Acompañado de preguntas relacionadas con aspectos sociodemográficos como identidad de género y rango de edad.

Los dos bloques siguientes del estudio consisten en idénticas baterías de preguntas, con escala Likert, donde se solicitaba señalar del 1 al 5 la elección que más se adecua de acuerdo con cada pregunta en atención primero a la fotografía y a continuación acerca de la ilustración. Las preguntas realizadas fueron:

- (P1). Equivalencia figurativa Literal: ¿La imagen representa la composición personajes y escenarios de la fiesta?
- (P2) Equivalencia figurativa Simbólica: ¿Representa la fiesta y su construcción mental?
- (P3) Gama Cromática: ¿Consideras que armoniza la luz y el color?
- (P4) Nivel de Familiaridad: ¿Te da más sensación de afinidad o confort?
- (P5) Nivel de Parecido Humano: ¿Crees que se parece más a la realidad?
- (P6) Significado Denotativo ¿Lo consideras: objetivo, literal, informativo?
- (P7) Significado Connotativo ¿Lo consideras: subjetivo, simbólico, emocional?
- (P8) Gusto, opinión: ¿Te gusta en rasgos generales?

En estas escalas Likert de 5 valores que van de (1 Nada; 2 Poco; 3 Moderado; Mucho; 5 Totalmente) se realizan idénticas preguntas, pero valorando en un primer bloque la Fotografía y en segundo bloque la ilustración generada.

3. Análisis de resultados

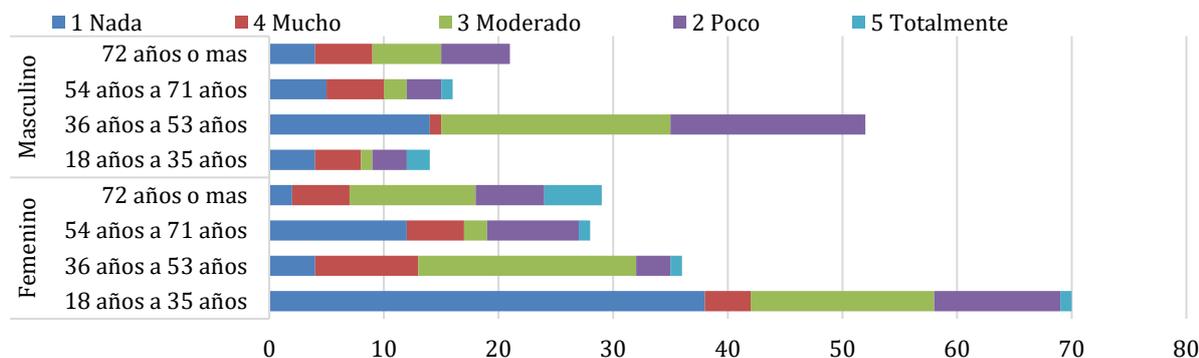
Las fiestas analizadas son 133 distintas, de las cuales el 8,27% tienen el máximo reconocimiento de internacional, el 13,53% el nacional y el 78,2% restante tan sólo regional. Cada ficha fue suministrada en dos ocasiones a un total de 266 encuestados distintos.

3.1. Conocimiento de la Fiesta

Como características sociodemográficas se les preguntó su rango de edad y sexo. Más de la mitad del muestreo (61,12%) fue femenino, de los cuales dos tercios eran menores de 53 años.

Como se observa en la figura 4, casi tercera parte de los encuestados (31%), afirmaron desconocer completamente la festividad que les tocó. Apenas el 22% confesó conocer la localidad donde se sitúa, pero desconocer la fiesta en sí. Poco más de la cuarta parte (29%) sabían de qué fiesta se trataba, pero no habían acudido en ninguna ocasión. Un 14% afirmaron conocer y haber estado en la fiesta en, al menos, una ocasión. Tan solo un 4% de los encuestado afirmó acudir todos los años y participar en la fiesta.

Figura 4. Respuestas a la pregunta: ¿Conocías o has estado en la Fiesta? Conocimiento de los encuestados por género y rango de edad.



Fuente: Elaboración propia, 2024.

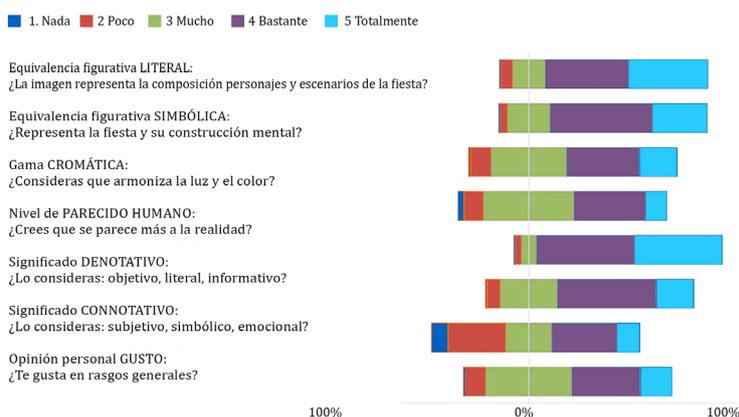
3.2. Valoraciones de los encuestados ante la fotografía real.

Se preguntó a los encuestados acerca de la fotografía situada en el lateral izquierdo que señalaran en una escala Likert (1 Nada; 2 Poco; 3 Moderado; 4 Mucho; 5 Totalmente). La característica donde la fotografía obtuvo más consenso es el Parecido Humano con una puntuación promedio más alta de 4,27. En la figura 5, observamos que ante la pregunta ¿Crees que se parece más a la realidad? es en lo que más de acuerdo están con un 89% que afirma que mucho o totalmente. En cambio, el aspecto de Gama Cromática, al ser preguntados si «¿Consideras que armoniza la luz y el color? Mientras que el Nivel de Familiaridad: ¿Te da más sensación de afinidad o confort?, son los que peor valoración ofrecen con un 10% de Poca adecuación»

El análisis de las preguntas ¿Lo consideras objetivo, literal, informativo? y ¿Lo consideras subjetivo, simbólico, emocional? Muestra claras diferencias. El Significado Denotativo es valorado positivamente por los encuestados donde casi la mitad 47,4% lo valoran como «Mucho» y casi la quinta parte 18,2% como «Totalmente». Siendo claramente inferior la valoración del Significado Connotativo donde el 27,7% lo considera como «Poco» y un 8% como «Nada».

Al ser preguntados sobre su opinión: ¿Te gusta en rasgos generales? Las respuestas son muy diversas, pero en general neutras con ligera tendencia positivas: «Poco» (16,8%), «Moderado» (28,5%) y «Mucho» (31,4%) aparecen más a menudo. Mientras que los extremos polarizados «Totalmente» y «Nada» no llegan a un cuarto (23,3%).

Figura 5. Resultados de fotografía, (Likert: 1 nada a 5 Totalmente).



Fuente: Elaboración propia, 2024.

3.3. Valoraciones de los encuestados ante la Ilustración generada con IA

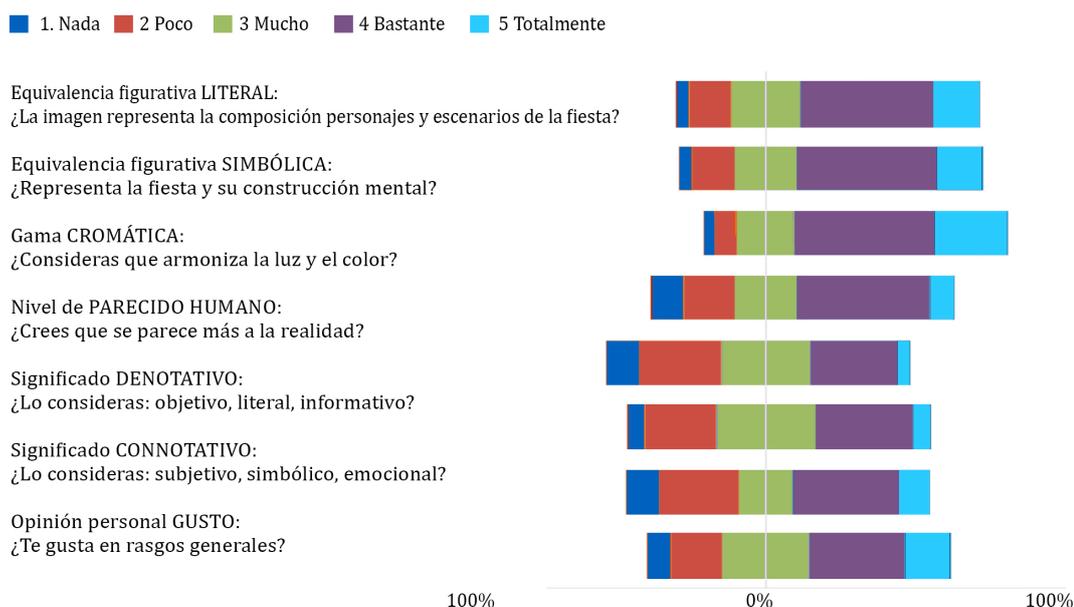
Se realizaron las mismas ocho preguntas, pero, en esta ocasión, acerca de la ilustración generada por IA. Los resultados muestran una frecuencia de aparición de «mucho» en todos los casos con rangos que van desde el 46% de Equivalencia figurativa Simbólica al 43,8% de Equivalencia Literal (Figura 6).

Los encuestados destacaron muy positivamente la Gama Cromática al considerar que la luz y el color armonizan «Mucho», con un 46%, o «Totalmente», con un 24,1%. En el lado contrario se encuentra el nivel de Parecido Humano, donde el 27% cree que se parece «Poco» y el 10,9% «Nada» a la realidad. Dato relacionado con el Nivel de Familiaridad donde, aunque casi la mitad (43,8%) considera «mucho» la sensación de afinidad o confort, un cuarto de los encuestados lo consideran «Poco» (16,8%) o «Nada» (10,9%) familiar.

Al ser preguntados sobre el Significado Denotativo y Connotativo, apenas se observan diferencias. Los encuestados ofrecen una media muy similar siendo ligeramente la percepción de «moderado» (32,8%) en cuanto a lo «objetivo, literal e informativo» de la imagen y algo más polarizado en cuanto el significado subjetivo, simbólico, emocional donde un (26,3%) lo consideran «Poco» y la tercera parte «Bastante» (35%). Es decir, hay algo más de disensión en cuanto a Connotación. Aunque el promedio sea prácticamente parejo: (3,07/5) en denotativo y (3,08/5) de promedio connotativo.

La opinión sobre la imagen generada por IA es positiva y casi la mitad de los encuestados, un 46,24%, consideran que les gusta mucho o totalmente, en contraste del 25,19% que aseguraban que les había gustado poco o nada, en rasgos generales.

Figura 6. Resultados de Ilustración generada con IA, (Escala Likert: 1 nada a 5 Totalmente).

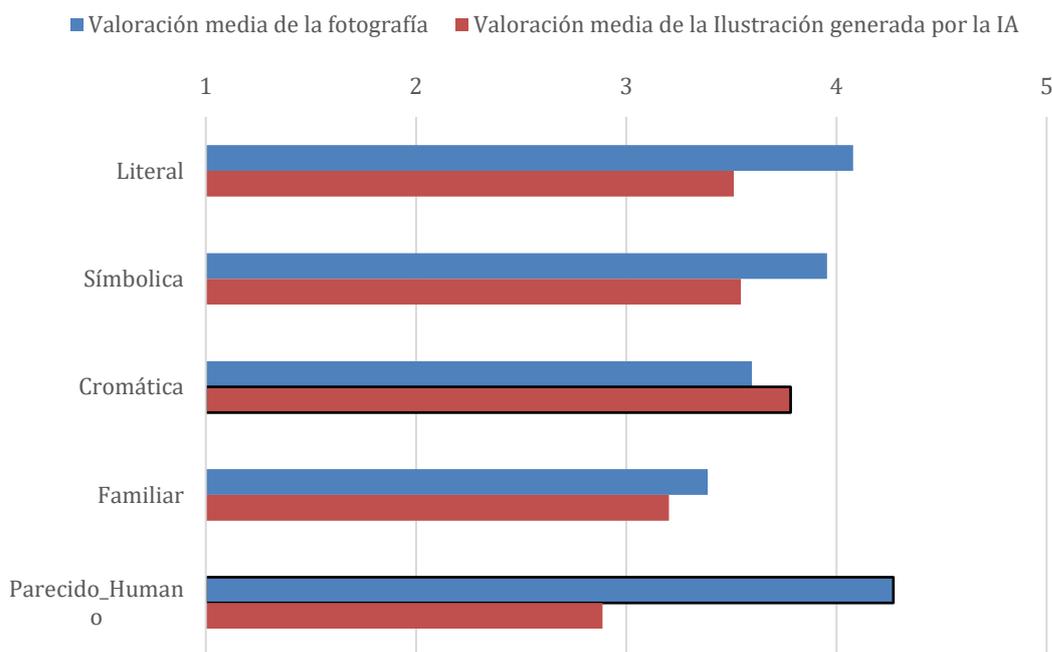


Fuente: Elaboración propia, 2024.

3.4. Cruce de datos

Si comparamos las respuestas, observamos en la figura 7 que la valoración media de la fotografía está determinada por la valoración de la imagen generada mediante IA. El valor que más contrasta es el Parecido Humano, donde se obtiene la máxima puntuación media (4,27/5) en la fotografía y la mínima (2,89/5) en la Ilustración. Existe una ligera diferencia de medio punto en los epígrafes Literal y Simbólica entre la valoración de fotografías e ilustraciones. En general, se considera mejor la composición, los personajes y los escenarios de las fotografías. Observamos la misma desviación en la representación de la fiesta y su construcción mental.

Figura 7. Resultados promedio de la valoración Literal, Simbólica, Cromática y Parecido Humano. (Escala Likert: 1 nada a 5 Totalmente).

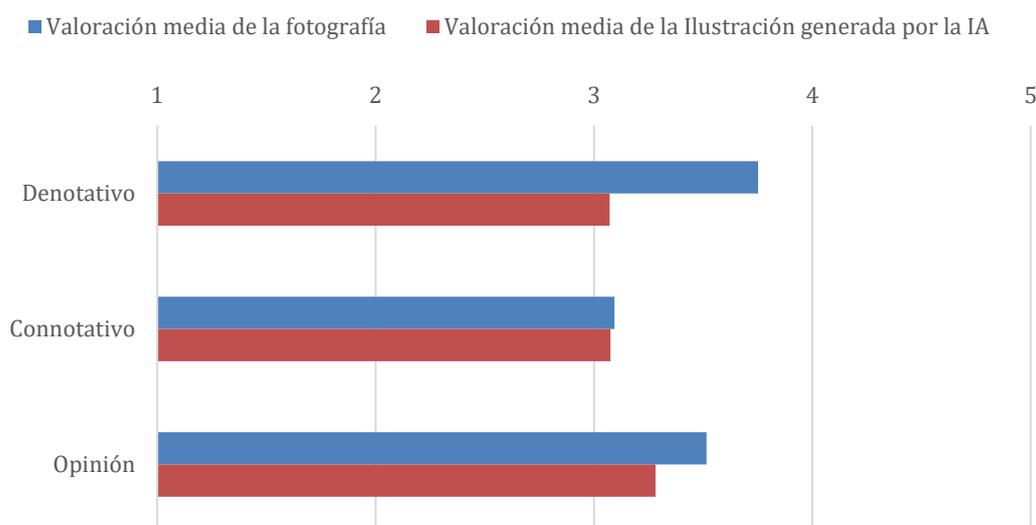


Fuente: Elaboración propia, 2024.

El único epígrafe en el que los encuestados han valorado mejor la imagen generada mediante IA es al preguntarles si consideran la armonización de luz y color. Este aspecto, que hemos denominado Gama Cromática, ofrece una ligera ventaja de 0,18 décimas a la imagen de IA.

Es sorprendente la poca diferencia en el nivel de confort. Cuando se les preguntó si les daba sensación de afinidad o confort los encuestados mostraron una ligera preferencia de 3,39/5 puntos hacia la fotografía, frente a los 3,2/5 puntos obtenidos por la ilustración.

Figura 8. Resultados promedio de la valoración Denotativa, Connotativa y Opinión. (Escala Likert: 1 nada a 5 Totalmente).



Fuente: Elaboración propia, 2024.

El significado Denotativo presenta una valoración muy distinta en la fotografía que en la ilustración generada (Figura 8). La diferencia de 0,68 puntos se debe a los valores Literal, Familiar y especialmente al de Parecido Humano. Estos valores implican consideraciones objetivas e informativas que explican esta diferencia.

En cambio, la Connotación refiere a valores simbólicos y emocionales. Estos aspectos están relacionados con las preguntas sobre equivalencia figurativa Simbólica, Cromática y la sensación subjetiva de familiaridad. Que no sólo han sido valorados con menor diferencia. Esto valida que la connotación obtenga, según los encuestados, un valor medio prácticamente idéntico, con una diferencia de menos de dos centésimas a favor de la fotografía.

4. Discusión y Conclusiones

La primera conclusión del estudio es el amplio desconocimiento y falta de percepción de las festividades analizadas entre los encuestados. En primer lugar, se observó un nivel significativo de desconocimiento general sobre las festividades, con la mayoría de los participantes admitiendo no tener conocimiento completo de la festividad asignada. Este hallazgo sugiere una falta de familiaridad o interés generalizado en las festividades estudiadas, lo que podría atribuirse a una variedad de factores, como la ubicación geográfica de los encuestados o la falta de promoción de estas festividades a nivel regional.

El proceso de evaluación de varias aplicaciones artificiales para transformar una imagen en una representación cultural válida revela la complejidad y los desafíos inherentes a la traducción de imágenes visuales a través de IA. Aunque se exploraron diversas opciones, ninguna plataforma logró una transformación completamente satisfactoria que conservara fielmente los elementos originales de la cultura y tradición representada en las fiestas. Si bien algunas herramientas como Leonardo AI mostraron un potencial prometedor, la necesidad de un entrenamiento adicional, sugiere que la mejora continua de estos sistemas es crucial para alcanzar resultados más precisos y fieles a las imágenes de referencia en futuras investigaciones y aplicaciones prácticas en el ámbito artístico y creativo.

En cuanto a la valoración de las imágenes por los públicos, tanto la fotografía real como la ilustración generada por IA recibieron una valoración generalmente positiva por parte de los encuestados. La fotografía real destaca por su similitud literal (P1) más que la simbólica (P2) con la realidad y en términos de parecido humano, aunque aspectos como la gama cromática (P3) y la sensación de afinidad y familiaridad (P4) obtuvieron calificaciones ligeramente más bajas. Por otro lado, la ilustración generada por IA también recibió valoraciones positivas, especialmente en aspectos como la luz (P3) y el color, aunque el parecido humano fue menos convincente para los encuestados. Estos hallazgos indican la efectividad tanto de la fotografía tradicional como de las representaciones generadas por IA para transmitir información visualmente, aunque cada una tiene sus propias fortalezas y debilidades en la percepción del público.

Uno de los aspectos más interesantes del estudio fue la influencia recíproca observada entre la valoración de la fotografía y la ilustración generada por IA. Se encontraron diferencias significativas en la percepción del parecido humano (P5) y la gama cromática (P3) entre ambas imágenes, lo que sugiere que la valoración de una imagen puede afectar la percepción de la otra. Esta interacción entre las dos formas de representación visual destaca la complejidad de la percepción humana y la importancia de considerar múltiples variables al evaluar imágenes generadas por inteligencia artificial.

Además, se identificaron diferencias significativas en la percepción del significado denotativo (P6) y connotativo (P7) entre la fotografía y la ilustración generada por IA. Mientras que la fotografía fue mejor valorada en aspectos objetivos e informativos (P5), la ilustración generada por IA recibió una valoración más favorable en términos de aspectos simbólicos y emocionales (P7). Estos hallazgos sugieren que la elección entre diferentes formas de representación visual puede influir en la interpretación y comprensión de un mensaje por parte del público, lo que tiene implicaciones importantes para la comunicación efectiva en diversos contextos.

En general, a pesar de las diferencias observadas en la percepción de las imágenes, tanto la fotografía como la ilustración generada por IA recibieron una opinión general positiva (P8) por parte de los encuestados. Esto sugiere un alto grado de aceptación y aprecio por parte del público hacia diferentes formas de representación visual, lo que respalda la idea de que la diversidad en la presentación de información puede enriquecer la experiencia del usuario y mejorar la comunicación efectiva.

El entrenamiento exhaustivo de Playground AI con la generación de más de 1000 imágenes de fiestas de interés turístico de Castilla y León revela incertidumbre sobre los límites del entrenamiento y enseñanza de inteligencias artificiales en este contexto cultural. A pesar de la extensiva interacción con la IA, no se observó una mejora significativa en la calidad, ni denotativa ni connotativa, de las imágenes generadas, tanto en inglés como en español, lo que sugiere que el proceso de entrenamiento puede no ser lineal o puede estar sujeto a limitaciones intrínsecas del sistema. Además, la variabilidad en el estilo de las imágenes creadas con el mismo *prompt* pone de relieve las restricciones predefinidas dentro de los algoritmos, lo que limita la capacidad de producir resultados más allá de ciertos parámetros establecidos. Estas limitaciones resaltan la necesidad de investigaciones adicionales para comprender mejor las capacidades y limitaciones de las IA en la creación de imágenes artísticas, así como para desarrollar estrategias efectivas de entrenamiento que fomenten la diversidad y la calidad de las producciones generadas.

La convergencia entre la IA y la creación artística ha sido objeto de un intenso escrutinio en los últimos años. Esta sinergia plantea preguntas fundamentales sobre la naturaleza misma del arte y la creatividad, así como sobre el papel del ser humano en el proceso creativo. En el contexto específico de esta investigación, la aplicación de la IA para generar imágenes no realistas, pero con relación con la realidad, presenta un desafío interesante por el gran número de posibilidades futuras tanto para el ámbito artístico como académico o investigador.

Desde su surgimiento, la IA ha demostrado una capacidad sorprendente para aprender y replicar patrones complejos en datos, incluidas imágenes visuales. Los avances en técnicas de aprendizaje profundo han permitido que los modelos de IA generen contenido visual que a menudo es indistinguible de las obras humanas, como lo demuestra la proliferación de *deepfakes* o de *fake news* que pueblan los medios digitales en los últimos años. Sin embargo, el enfoque de esta investigación va más allá de simplemente imitar la realidad; ha buscado explorar las posibilidades creativas inherentes a la IA como herramienta artística.

La elección de utilizar herramientas no realistas en la generación de imágenes a través de la IA es significativa en varios aspectos. En primer lugar, desafía las convenciones tradicionales de representación visual al subvertir las expectativas de realismo. Esta ruptura con la norma invita a reflexionar sobre la naturaleza misma de la imagen y su relación con la realidad percibida, haciéndonos pensar en que significa ser fiel a la realidad en un mundo donde la línea entre lo real y lo artificial se vuelve cada vez más difusa.

Además, esta investigación plantea preguntas sobre la naturaleza de la creatividad y la autoría en el contexto de la IA, como si es posible considerar que una obra generada por IA es realmente "creativa" en el sentido humano de la palabra, o qué papel juega el artista en el proceso de creación cuando utiliza herramientas computacionales para generar arte. Los investigadores participantes del proyecto en ningún momento se han sentido en el rol de creadores o artistas audiovisuales digitales; por lo tanto, este tipo de definiciones se entienden como etiquetas con una mayor relación con el futuro mercado audiovisual, un término que actualmente domina los mercados digitales y que hace años no era actividad relacionada con la vida laboral.

Por otro lado, la decisión de evitar la generación de imágenes excesivamente realistas también plantea desafíos técnicos y estéticos. El equilibrio entre la representación fiel de la realidad y la expresión artística puede resultar complicado, especialmente cuando se utilizan algoritmos de IA diseñados para maximizar la precisión y la fidelidad visual. En este sentido, la investigación no solo se enfrenta a cuestiones conceptuales, sino también a desafíos prácticos en la implementación de su enfoque.

A raíz de los resultados de este estudio, observamos que las ilustraciones de IA analizadas no son tan bien valoradas en aspectos denotativos como las fotografías reales. Esto nos revela nuestro segundo objetivo específico (OE2) indicando que la retórica visual de Barthes se integra en la generación de imágenes de manera desigual. También los resultados señalan (OE2) que logran sortear el Valle Inquietante mediante la inclusión de recursos técnicos y estéticos. Esto nos permite abordar el objetivo principal, es decir, que la teoría del Valle inquietante de Masahiro Mori también se aplica a las imágenes generadas por inteligencia artificial. Es por esto que Playground AI sortea la inclusión de representación de imágenes de alto parecido humano, aunque no siempre con éxito.

En última instancia, el enfoque adoptado en esta investigación abre nuevas vías para la exploración artística y tecnológica. Al desafiar las convenciones establecidas y experimentar con nuevas formas de

expresión visual, se abre un espacio fértil para la innovación y el descubrimiento. Más allá de los resultados específicos de esta investigación, su impacto radica en su capacidad para inspirar nuevas preguntas y perspectivas sobre el arte, la tecnología y la intersección entre ambos.

Destacamos la importancia de considerar tanto aspectos objetivos como subjetivos al evaluar imágenes en contextos culturales. Además, subrayan el potencial de las tecnologías generativas de imágenes mediante inteligencia artificial para generar representaciones visuales convincentes. Señalamos la necesidad de investigaciones adicionales para comprender mejor las complejidades de la percepción humana y su impacto en la comunicación en cultura visual. Estas conclusiones proporcionan una base sólida para futuras investigaciones en este campo y tienen implicaciones importantes para el diseño y la presentación de información en una variedad de contextos culturales.

5. Agradecimientos

Este proyecto ha sido financiado por la Junta de Castilla y León bajo el proyecto: Servicio de elaboración de Fiestas de Interés Turístico (Internacional, Nacional y de Castilla y León) en la Comunidad de Castilla y León. Expediente: B2023/018077.

Referencias

- Allen, G. (2011). *Intertextuality*. Routledge.
- Bravo Heredia, A. (2010). Peirce: hacia una semiótica visual transdisciplinaria. Actas de las segundas Jornadas internacionales Peirceanas. *Educación, arte y signo Actas de las Segundas Jornadas Internacionales Peirceanas*, 77-87.
- Cwaik, J. (2020). *7R: Las siete revoluciones tecnológicas que transformarán nuestra vida*. Ed. Conecta.
- Estupiñán Ricardo, J., Leyva Vázquez, M. Y., Peñafiel Palacios, A. J. y El Assafiri Ojeda, Y. (2021). Inteligencia artificial y propiedad intelectual. *Universidad y Sociedad*, 13(S3), 362-368. Recuperado a partir de <https://rus.ucf.edu.cu/index.php/rus/article/view/2490>
- Gómez de Ágreda, A., Feijoo González, C. A. y Salazar García, I. A. (2021). Una nueva taxonomía del uso de la imagen en la conformación interesada del relato digital. Deep fakes e inteligencia artificial. *El Profesional de la Información*, 30 (2); 1-24. ISSN 1386-6710. <https://doi.org/10.3145/epi.2021.mar.16>
- Gómez Gómez, H. y Rubio Tamayo, J. L. (2023). Algorithmgraphy as a milestone and phenomenon in the production of still images in the digital era: Resignification of the notion of the photographic image and projection of the medium in a context of image production with artificial intelligence and machine learning. *Visual Review. International Visual Culture Review Revista Internacional De Cultura Visual*, 14(2), 1-13. <https://doi.org/10.37467/revvisual.v10.4607>
- Heras Benavides, D.A. (2017). Clasificador de imágenes de frutas basado en inteligencia artificial. *Revista Killkana Técnica*, 1(2), 21. https://doi.org/10.26871/killkana_tecnica.v1i2.79
- Icarte Ahumada, G. A. (2016). Aplicaciones de inteligencia artificial en procesos de cadenas de suministros: una revisión sistemática. *Ingeniare. Revista chilena de ingeniería*, 24(4), 663-679. <http://dx.doi.org/10.4067/S0718-33052016000400011>
- Jordanous, A. (2012). A Standardised Procedure for Evaluating Creative Systems: Computational Creativity Evaluation Based on What it is to be Creative. *Cognitive Computation*, 4(3), 246-279. <https://doi.org/10.1007/s12559-012-9156-1>
- Lin, H. (2018). Developing responses to cyber-enabled information warfare and influence operations. *Lawfare*, September 6. <https://tinyurl.com/228hj989>
- Lizama, F. (2024). ChatGPT: la adopción disruptiva de la inteligencia artificial. *Revista Economía & administración Mirada FEN*, No. 184. 46-48. <https://mirada.fen.uchile.cl/uploads/revista/Mirada-FEN-184/66/index.html>
- López Cardona, A., Molina Saldarriaga, M.L. y Serna Usme, D. (2020). *Composición visual: Análisis de variables en la construcción de imágenes*. Institución Universitaria de Envigado.
- McCormack, J., Hutchings, P., Gifford, T., Yee-King, M., Llano, M. T. & D'inverno, M. (2020). Design Considerations for Real-Time Collaboration with Creative Artificial Intelligence. *Organised Sound*, 25(1), 41-52. <https://doi.org/10.1017/S1355771819000451>
- Menéndez-Pidal, S. N. (2010). Retórica visual: una herramienta necesaria en la creación e interpretación de productos visuales. *Revista de artes y humanidades Única*, 11(2), 99-116. <https://www.redalyc.org/pdf/1701/170121899006.pdf>
- Mengoni, A. (2021). Visual Semiotics. In: Purgar, K. (eds) *The Palgrave Handbook of Image Studies*. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-71830-5_39
- Milo, A. (2023). Uncanny valley: qué es la teoría del valle inquietante y cómo explica nuestro miedo a los robots. *National Geographic*. <https://tinyurl.com/2ysrte8t>
- Montaña, R. de la (2001). *Pingada del Mayo* [Fotografía]. Página web del ayuntamiento de Hontoria del Pinar. <https://www.hontoriadelpinar.es/fiestas-y-tradiciones/pingada-del-mayo>
- Playground AI. (2023). Playground (versión septiembre 2023) [plataforma de IA]. <https://playground.com/>
- Toca Torres, C. E. (2014). Inteligencia colectiva: enfoque para el análisis de redes. *Estudios Gerenciales*, 30(132), 259-266. <https://doi.org/10.1016/j.estger.2014.01.014>
- Tostado Sánchez, S. E., Ornelas Rodríguez, M., Espinal Jiménez, A. y Puga Soberanes, H. J. (2016). Implementación de Algoritmos de Inteligencia Artificial para el Entrenamiento de Redes Neuronales de Segunda Generación. *Jóvenes en la ciencia*, 2(1), 6-10. <https://www.jovenesenlaciencia.ugto.mx/index.php/jovenesenlaciencia/article/view/715>
- Villafañe, J. y Mínguez, N. (2002), *Principios de teoría general de la imagen*, Madrid, Ediciones Pirámide.

Yoguel, G., Chanders, V. y Mochi, S. (2021). Innovación por coproducción en industria 4.0: un estudio de caso de inteligencia artificial aplicadas a imágenes. Ciudad Autónoma de Buenos Aires: *CIECTI*.
<http://www.ciecti.org.ar/wpcontent/uploads/2021/04/DT23-V05.pdf>