

El uso de la inteligencia artificial en la reconstrucción de eventos artístico-educativos

The use of artificial intelligence in the reconstruction of artistic and educational events

Ricardo Campos García

Investigador independiente

ricardocampos@correo.ugr.es

Recibido: 02/02/2024

Revisado: 03/05/2024

Aceptado: 23/06/2024

Publicado: 01/07/2024

Sugerencias para citar este artículo:

Campos García, Ricardo (2024). «El uso de la inteligencia artificial en la reconstrucción de eventos artístico-educativos», *Tercio Creciente*, 26, (pp. 83-93), <https://dx.doi.org/10.17561/rtc.26.8660>

Resumen

El incremento en la aparición y en el uso de aplicaciones de inteligencia artificial, está produciendo una creación masiva de elementos visuales de características muy similares. Pero, al mismo tiempo, esta tecnología ofrece la posibilidad de crear mundos fascinantes que de otra manera sería, o bien imposible, o muy costoso a nivel presupuesto y tiempo.

Para este artículo se plantea un uso alternativo que consiste en aprender la historia de nuestro arte, pero a través de la creación de ficciones visuales, es decir, de la composición de falsas realidades que nos ayuden a entender supuestos que nunca ocurrieron, pero que pudieron llegar a suceder.

Palabras clave: inteligencia artificial, arte, artes visuales, IA, ficción visual.

Abstract

The increasing emergence and use of artificial intelligence applications is leading to a massive creation of visual elements with very similar characteristics. But, at the same time, this technology offers the possibility of creating fascinating worlds that would otherwise be either impossible or very costly in terms of budget and time.

For this article we propose an alternative use that consists of learning the history of our art, but through the creation of visual fictions, that is, the composition of false realities that help us to understand assumptions that never happened, but that could have happened.

Keywords: Artificial Intelligence, Art, Visual Art, AI, Visual fiction.

1. Introducción

En esta investigación se plantea el uso de las herramientas de inteligencia artificial para crear una ficción académica. Este propósito se lleva a cabo realizando una recreación digital de eventos visuales, inspirados en el estilo de dos artistas que fueron coetáneos, pero que nunca llegaron a conocerse: Dorothea Lange y Juan Gris.

Dorothea Lange (1895 - 1965) fue una fotoperiodista documental, muy conocida por sus crónicas en la llamada Gran Depresión en Norteamérica. De sus miles de obras destacan las que realizó sobre las condiciones de vida de los trabajadores y sus familias, en las granjas del sur y del oeste de Estados Unidos en la década de los años 30 del siglo XX (Lange y Taylor, 1939). Sus imágenes son un espejo de quiénes éramos, cómo llegamos a ser y en qué estábamos en proceso de convertirnos. El estilo de Dorothea se identifica por la manera que tiene de componer retratos y escenas donde casi siempre, una figura humana es parte de la arquitectura del espacio (Whinston Spirn, 2008).

Juan Gris (1887 - 1927) fue ilustrador y pintor, considerado uno de los grandes maestros del cubismo. Sus pinturas, aunque menos numerosas que las de algunos de sus grandes contemporáneos, son consideradas de tan alta calidad que le definen como un artista original y único dentro del vocabulario del cubismo. Según James T. Soby (1958), Juan Gris se interesó especialmente por los efectos atmosféricos y las formas de la compleja naturaleza muerta. Le interesaban, más que a cualquiera de sus colegas en el Cubismo, las modulaciones físicas del entorno en el que trabajaba.

Es un reto utilizar la inteligencia artificial para recrear escenas o situaciones históricas recientes, desarrollando el concepto de “docuficción”. La docuficción, permite seguir con la tradición del juego metaficcional, y al mismo tiempo permite ocuparse de temas relevantes para la sociedad y la política (Martínez Rubio, 2014). Igualmente se puede centrar en asuntos que refieren a ficciones artísticas o culturales. A través de elementos ficticios podemos obtener una explicación de lo real. En definitiva, se trata de un nuevo campo de producción y figuración de la verdad en las imágenes visuales.

La inteligencia artificial permite realizar un entrenamiento de estilos. Aunque no es un uso frecuente, a través de plataformas de software libres es posible modificar y reforzar los modelos utilizados a nivel general. Estos modelos han sido entrenados con decenas de millones de imágenes a través de complejos procesos algorítmicos, pero es posible modificarlos para que se adapten mejor a nuestras propias intenciones visuales y conceptuales. Llevar a cabo estos entrenamientos no es tan fácil como usar plataformas en línea como Dall-E o Midjourney (las dos más populares y de uso más mayoritario). Stable Diffusion es un ejemplo de plataforma de inteligencia artificial que permite este tipo de entrenamientos a partir de un grupo de imágenes que hayamos seleccionado. La herramienta fue desarrollada por la empresa Stability AI, empresa norteamericana que comenzó su andadura en 2019, intentando “democratizar el uso de la Inteligencia Artificial construyendo una base global para activar el potencial de la humanidad” (Stability AI, 2019, s. p.).

Existen varias formas de realizar un entrenamiento de estilos, al igual que se pueden crear procesamientos de figuras humanas y de personas individuales, rostros, animales u objetos concretos. Lo importante es realizar una acertada selección de contenidos visuales (un grupo de imágenes en formato PNG) a las que se adjunta una descripción escrita (en formato *.txt) que describa lo que el sistema debe interpretar sobre lo que está analizando en cada imagen.

En el año 2023, Coca-Cola (2023) creó un anuncio comercial recreando ficciones visuales, en el que se integraban artistas clásicos con elementos contemporáneos. El uso de modelos generales que son complementados con entrenamientos de estilos específicos es algo que se mantiene discretamente silenciado en producciones artísticas y audiovisuales. Exceptuando aquellos artistas que se explicita y abiertamente se postulan como creadores de contenido a través de inteligencias artificiales generativas, o perfiles que comparten sus creaciones personales de forma anónima a través de la plataforma de mensajería Reddit (reddit.com), es bastante complicado localizar proyectos donde este tipo de entrenamientos se utilicen de manera profesional.

Por otro lado, la mayoría de aplicaciones de inteligencia artificial de fácil acceso (donde podemos elegir efectos, filtros o determinados estilos) han sido creadas previamente con entrenamientos de estilos en aplicaciones como Stable Diffusion para, a continuación, crearles una máscara o entorno cómodo de usar y ponerlo a disposición del público. Es el caso de páginas como Dream Studio o Play Ground.

2. Objetivos

En primer lugar. Dilucidar o comprobar la credibilidad o veracidad de las ficciones visuales creadas por la inteligencia artificial.

Segundo, encontrar nuevos usos de la inteligencia artificial en la creación de imágenes, los cuales pueden ser de utilidad en planteamientos artísticos o en producciones audiovisuales.

Tercero. Determinar sus aplicaciones educativas. ¿Qué se aprende con este tipo de recreaciones o ficciones visuales? ¿Son únicamente un elemento de diversión o pueden ofrecer un contenido cultural y educativo?

Hasta hace apenas unos pocos meses, construir este tipo de ficciones visuales suponía un reto complejo y costoso, lejos del alcance para cualquier persona que no fuese profesional o especialista. La actual revolución tecnológica en torno a la creación de contenido visual ha producido un cambio de escenario, facilitando enormemente la producción de contenidos visuales con la inteligencia artificial generativa.

3. Proceso

¿Cómo era un taller con alumnado y estudiantes en el París de los años 1920, en la órbita del cubismo? ¿Qué fotografías tendríamos disponibles si la fotógrafa norteamericana Dorothea Lange hubiese coincidido con el pintor Juan Gris en su taller?

Para conseguir estas imágenes, se plantean tres procesos de trabajo.

Primero. Generar imágenes que simulen que han sido creadas desde el propio ojo de Dorothea Lange usando directamente plataformas de IA. En este punto se utiliza Dall-E, una aplicación en línea que se puede ejecutar a través de Chat-GPT4 (en su versión más moderna) o a través de la página de sus creadores, OpenAI (para versiones anteriores y de menor calidad visual). En este caso he utilizado la versión más reciente hasta la fecha. Es importante destacar que para este primer proceso no se realiza ningún entrenamiento específico de estilos, sino que se utiliza el contenido que está disponible para cualquier usuario.

Segundo, y más importante, llevar a cabo un procesamiento específico de un conjunto de datos [dataset] compuesto por fotografías originales de Dorothea Lange que se acompañan de sucintas descripciones en texto.

En tercer lugar, realizar una comparación de resultados entre el primer y el segundo proceso.

3.1. Creación visual con Dall-E.

A la hora de solicitar el contenido a Dall-E, a través de la aplicación de ChatGPT 4, se obtiene un bloqueo restrictivo por parte de la propia herramienta el cuál se muestra textualmente.

-Investigador: Me gustaría generar imágenes en blanco y negro, con el estilo de Dorothea Lange, pero simulando que la artista está fotografiando el estudio del pintor Juan Gris, en el París de 1920. En este estudio está el propio Juan Gris con sus discípulos.

-ChatGPT: *Lamento informarte que no puedo generar imágenes que imiten directamente el estilo de artistas específicos, como Dorothea Lange, debido a nuestra política de contenido. Sin embargo, puedo crear una imagen que capture la esencia de un estudio de arte en el París de 1920, con un enfoque en el estilo cubista de Juan Gris y un ambiente inspirado en las fotografías en blanco y negro de la época. ¿Te gustaría que proceda con esta idea?*

-Investigador: Pero el estilo de Juan Gris y Dorothea Lange ya forman parte de la cultura general y además lo quiero utilizar para un estudio académico...

-ChatGPT: *Entiendo tu punto de vista y agradezco tu interés en el estilo de Dorothea Lange y Juan Gris. A pesar de que sus obras son ampliamente reconocidas y forman parte de la cultura general, nuestra política de contenido en la creación de imágenes se enfoca en evitar la imitación directa de estilos de artistas específicos, especialmente si sus trabajos son relativamente modernos y pueden estar sujetos a derechos de autor o a consideraciones de propiedad intelectual. Sin embargo, puedo ayudarte a crear una imagen inspirada en la época del París de los años 1920, con un enfoque en el arte abstracto y la fotografía en blanco y negro, sin atribuir directamente el estilo a un artista específico. Si esto te parece bien, por favor házmelo saber y procederé con la creación de una imagen con estas características generales.*

- Investigador: Sí, por favor.

-ChatGPT: *Aquí tienes la imagen que evoca la atmósfera artística del París de los años 1920 en un estudio de arte. Esta imagen en blanco y negro muestra un espacio lleno de actividad artística, con artistas creando y discutiendo obras de arte abstracto, rodeados de pinturas y esculturas que reflejan las tendencias abstractas y geométricas del arte de principios del siglo XX. Espero que capture la esencia de la comunidad artística histórica que deseabas.*



Figura 1. Autor + IA (2023). *Estudio de arte en París en los años 20 del siglo XX*. Dall0001.
Imagen digital.

3.2. Creación visual con Stable Diffusion, versión 1.5.

En la plataforma Stable Diffusion, con los parámetros por defecto de la aplicación, introduce la siguiente instrucción escrita (en inglés): “Fotografía del estudio del artista cubista Juan Gris, con sus discípulos, en París en 1920. La imagen está tomada por Dorothea Lange, por lo que tiene su propio estilo. Los artistas están trabajando. Buenas caras. Caras perfectas.”



Figura 2. Autor + IA (2023). *Estudio de arte en París en los años 20 del siglo XX*.SD0001.
Imagen digital.

Como puede apreciarse en la Figura 2, la definición de los rostros, manos y otros elementos corporales son bastante mejorables, por lo que es necesario seguir avanzando con el entrenamiento específico de las plataformas de inteligencia artificial.

3.3. Creación visual con Stable Diffusion, versión XL. Modelo original.

Pasé a utilizar esta versión porque es la más preparada para la generación de imágenes fotorrealistas, por haber sido entrenada con un mayor número de imágenes de este tipo.



Figura 3. Autor + IA (2023) *Una fotografía del taller cubista de Juan Gris como profesor y algunos estudiantes de arte haciendo dibujos, en París a principios del siglo XX, estilo Dorothea Lange.* Imagen SD0006 Imagen digital.

La versión XL de Stable Diffusion no solo ofrece mejores rostros humanos, sino que los dibujos y pinturas que aparecen en cada una de las escenas simulan con mayor verosimilitud obras cubistas. Además, esta versión permite trabajar mejor la importancia de determinados conceptos dentro de las instrucciones escritas [*prompt*] para crear imágenes visuales, dándole mayor peso a un término, por ejemplo ‘realista’, que a otro término, por ejemplo ‘rostro humano’.



Figura 4. Autor+IA (2024) *Una fotografía del taller cubista de Juan Gris como profesor y algunos estudiantes de arte haciendo dibujos, en París a principios del siglo XX*, estilo Dorothea Lange. Imagen SD0010. Imagen digital.

3.4. Stable Diffusion, versión XL. Modelo pre-procesado.

Para llevar a cabo un entrenamiento personalizado de la inteligencia artificial con fotografías de Dorothea Lange, el primer paso fue seleccionar un grupo de 41 imágenes originales de la artista, que agrupé en dos bloques, uno más centrado en el reconocimiento del estilo de la fotógrafa y, otro en el aprendizaje de rostros y facciones.

Cada uno de los archivos de imagen, en formato 1024 por 1024 píxeles, lo acompañé de un archivo de texto con una descripción de la fotografía. Para esta descripción, hice uso de la herramienta “*Clip Interrogator*”, extensión disponible para instalar en la propia herramienta Stable Diffusion. Esta extensión asocia cada imagen con una descripción escrita. Usé la opción ‘rápido’ [*fast*] del modelo ‘ViT-L-14/openai’, uno de los muchos procesadores que tiene disponibles, y el más recomendado para el modelo XL de Stable Diffusion. A modo de ejemplo, transcribo dos de las descripciones verbales generadas por este procedimiento, traducidas al español:

- Dos niñas sentadas en un banco leyendo un libro, por Dorothea Lange, inspirada por Dorothea Lange, por George Herbert Baker, fotografía de Bill Brauer, por Mary Abbott, fotografía de archivo, fotografía vintage en blanco y negro, inspirada por Antanas Sutkus, por William Grant Stevenson, inspirada por Mary Abbott, fotografía de 1940. Dos niñas sentadas en un banco leyendo juntas un libro, una de ellas está mirando una revista. (((Estilo Dorothea Lange)))
- Un hombre sentado en un coche con un volante, por Dorothea Lange, wpa, inspirado por Dorothea Lange, fotografía de Bill Brauer, por Edward Weston, por Sergio Larraín, por Kati Horna, por Ben Shahn, foto de un hombre, fotografía de los años 30. un hombre sentado en el asiento del conductor de un coche con un volante y una manta en el suelo. (((estilo Dorothea Lange))).

Uno de los elementos más llamativos de estas descripciones escritas es la mención de nombres de pintores y pintoras, fotógrafos y fotógrafas, a veces compatriotas aproximadamente coetáneos de Dorothea Lange, pero en muchas ocasiones escasamente relacionados con ella ni por su época, ni por su estilo o temática. Es cierto que en ocasiones los artistas mencionados pueden suscitar analogías o correspondencias insólitas y bastante originales, pero sería difícil determinar si son más útiles o certeras que emparejamientos aleatorios.

El siguiente paso consiste en la ejecución de dos entrenamientos distintos.

En el primero se trabajan todas las imágenes seleccionadas. Este entrenamiento genera un archivo de bajo rango o *LORA*, que tiene un peso en el disco duro de 340 megabytes. *LORA* es el acrónimo en inglés de *Low Rank Adaptation* [adaptación de bajo rango] y permite entrenamientos más rápidos y con menos volumen de imágenes (Hu et al., 2021). Este módulo ha sido desarrollado por las empresas Microsoft y Simo, e integrado en Stable Diffusion.

El segundo entrenamiento está creado únicamente con algunas imágenes de retratos y rostros de Dorothea Lange, para mejorar las facciones de un rostro humano que, por defecto, suelen ser algo deformes y poco realistas.

Una vez terminado este procesamiento, para el que se dedican aproximadamente unas 3 horas, y se realizan un total de 1200 pasos de aprendizaje, ya se dispone de los dos archivos *LORA*, por lo que ya es posible utilizarlos en la creación de nuevas imágenes.

3.5. Afinamiento de los detalles de rostros humanos.

Para comprobar si todo el proceso se ha realizado correctamente, escribo una instrucción [*prompt*] muy sencilla, tal como ‘una fotografía de una mujer’, aplicándola al primer entrenamiento.



Figura 5. Autor + IA (2024) *Doble retrato*. Compuesto por dos imágenes digitales, izquierda *Imagen SD0018*, y derecha *Imagen SD0017*.

El estilo es correcto y los retratos son bastante realistas, aunque con pequeños errores que pueden parecer daños del paso del tiempo en el papel fotográfico. La petición de retratos, cercanos mirando a cámara, suele ofrecer mejores resultados que una escena más compleja, pero, en cualquier caso, el principal objetivo de esta solicitud es comprobar la calidad visual de los resultados tras los entrenamientos específicos.

Existe una segunda opción para mejorar aún más los retratos de personas, y es importando un modelo de Stable Diffusion XL entrenado por la comunidad de usuarios de la herramienta. En este caso, el modelo se llama 'RealVisXL V3.0 + Turbo' y se puede descargar libremente en la plataforma Civitai (civitai.com), donde los usuarios comparten sus entrenamientos y los ponen a disposición de la comunidad. Este modelo, aplicado a los entrenamientos específicos que había realizado produce imágenes sorprendentemente realistas y mantiene el estilo fotográfico buscado.

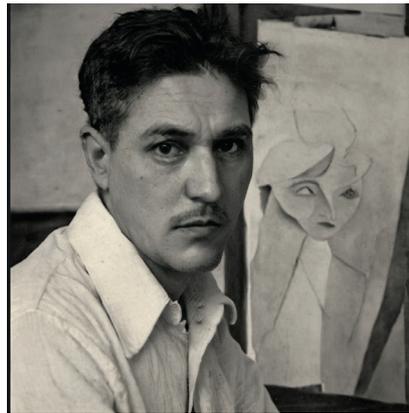


Figura 6. Autor + IA (2024) *Retrato de Juan Gris*. Imagen SD0019.

El siguiente paso consiste en solicitar imágenes centradas en el estudio de Juan Gris con sus discípulos.

4. El estudio de Juan Gris con alumnado

La redacción escrita de las instrucciones o peticiones [*prompt*] es necesario ir afinándola en cada entrenamiento, principalmente probando con diferentes prioridades de conceptos. Cuando más en el inicio del texto se encuentre un concepto, mayor importancia tendrá en el cálculo para la creación de la imagen.



Figura 7. Autor + IA (2024) *Una fotografía dentro del estudio de arte de Juan Gris con sus alumnos pintando <lora:DrtlngXL:1>, mirando directamente a la cámara, expresión segura y amable, atuendo elegante, fondo de escena urbana callejera. Imagen SD0030. Imagen digital.*



Figura 8. Autor + IA (2024) (<lora:DrtlngFCEsXL:1> style:1.2), *personas pintando con Juan Gris en su estudio, en París 1920, mirando a la cámara, concentrados, obras de arte cubistas. Imagen SD0036. Imagen digital.*

5. Conclusión

Trabajar con la inteligencia artificial para crear nuevas imágenes visuales no es tan sencillo ni tan azaroso como pedir un texto y recibir una imagen. El esfuerzo previo enfocado en el estudio, análisis, selección y procesamiento del contenido previo a un entrenamiento

es una labor de vital importancia. Preprocesar y entrenar es la manera más recomendable para poder trabajar en la creación de imágenes con inteligencia artificial visual al menos desde un sentido profesional enfocado a la enseñanza.

Las potencialidades de la inteligencia artificial para la creación de imágenes se encuentran en la variedad de opciones, combinaciones y extensiones que puede llegar a ofrecer. Tantas, que es fundamental disponer de un concepto artístico claro y bien definido.

Referencias

- Coca-Cola. (2023). *Masterpiece*. [Video]. YouTube. <https://www.youtube.com/watch?v=ZbYz-hg2zy4>
- Hu, E. J., Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; y Chen, W. (2021). LoRA: Low-Rank Adaptation of Large Language Models. *ArXiv*. <https://doi.org/10.48550/arXiv.2106.09685>
- Lange, D. and Taylor, P. S. (1939). *An American exodus: A record of human erosion*. Reynal & Hitchcock.
- Martínez Rubio, J. (2014). Autoficción y docuficción como propuestas de sentido. Razones culturales para la representación ambigua. *Castilla: Estudios de Literatura* (5):26-38. <https://revistas.uva.es/index.php/castilla/article/view/213/214>
- Soby, James T. (1958). *Juan Gris*. Museum of Modern Art, New York.
- Stability AI. (2019). About. *Stability AI*. <https://stability.ai/about>.
- Whinston Spirn, A. (2008). *Daring to Look: Dorothea Lange's Photographs and Reports from the Field*. University of Chicago Press.