

INTERACCIÓN HUMANO-MÁQUINA IMPULSADA POR INTELIGENCIA ARTIFICIAL PARA EL DISEÑO Y REPRESENTACIÓN DE ARQUITECTURA

AI DRIVEN HUMAN-MACHINE INTERACTION IN ARCHITECTURAL DESIGN AND REPRESENTATION

Federico Luis del Blanco García; orcid: 0000-0002-7907-6643

Sarah Cobos Martín; orcid: 0009-0001-2197-3149

UNIVERSIDAD POLITÉCNICA DE MADRID

doi: 10.4995/ega.2025.23147

Este artículo explora la interacción entre dibujos realizados a mano y modelos generativos de inteligencia artificial, destacando el control sobre los resultados para reducir la aleatoriedad en los procesos creativos. Mediante la implementación de un modelo de entrenamiento en Stable Diffusion, se logró establecer un estilo gráfico personalizado y consistente, garantizando coherencia y uniformidad en las representaciones visuales. Posteriormente, se aplicaron diferentes técnicas de representación a tres entornos, emulando un proceso arquitectónico que incluye bocetos, dibujos a lápiz, acuarelas, dibujos técnicos a tinta, modelado 3D y renderizado fotorrealista. Este enfoque muestra cómo la inteligencia artificial puede mejorar la representación de ideas arquitectónicas al facilitar un proceso creativo dinámico y colaborativo entre humanos y máquinas. Los resultados

destacan que el uso estratégico de la IA optimiza técnicas de representación, expande las posibilidades creativas y actualiza métodos tradicionales, manteniendo la influencia del diseñador y promoviendo flujos de trabajo innovadores en la arquitectura.

PALABRAS CLAVE: INTELIGENCIA ARTIFICIAL; STABLE DIFFUSION; DISEÑO GENERATIVO; DREAMBOOTH; LORA; REPRESENTACIÓN DE ARQUITECTURA; MIDJOURNEY; DALL-E; MODELOS DE DIFUSIÓN LATENTE

This article explores the interaction between hand-made drawings and generative artificial intelligence models, emphasizing control over the results to reduce randomness in the creative process. By implementing a training model in Stable Diffusion, a consistent and predictable graphic style was developed, ensuring uniformity

in the visual representations. Subsequently, various representation techniques were applied across three environments, emulating an architectural process that includes sketches, pencil drawings, watercolors, technical ink drawings, 3D modeling, and photorealistic rendering. This approach demonstrates how artificial intelligence can enhance the representation of architectural ideas by facilitating a dynamic and collaborative creative process between humans and machines. The results highlight that the strategic use of AI optimizes representation techniques, expands creative possibilities, and modernizes traditional methods, while maintaining the designer's influence and promoting innovative workflows in architecture.

KEYWORDS: ARTIFICIAL INTELLIGENCE; STABLE DIFFUSION; GENERATIVE DESIGN; DREAMBOOTH; LORA; ARCHITECTURAL REPRESENTATION; MIDJOURNEY; DALL-E; LATENT DIFFUSION MODELS



Introducción

La inteligencia artificial (IA) se ha convertido en una fuerza transformadora en diversos campos, y la arquitectura no es una excepción. La integración de la interacción humano-máquina impulsada por IA en el diseño y la representación arquitectónica ofrece oportunidades sin precedentes para mejorar la creatividad, la eficiencia y la precisión. Al aprovechar algoritmos avanzados y técnicas de aprendizaje automático, los arquitectos y diseñadores pueden explorar enfoques innovadores para el diseño, conceptualización o visualización de arquitectura.

A medida que la IA continúa evolucionando, su papel en la arquitectura está destinado a crecer, transformando la manera en que concebimos y comunicamos las ideas (Carpo 2023). Desde la generación de geometrías complejas hasta la optimización de soluciones sostenibles, los sistemas impulsados por IA permiten a los arquitectos superar los límites de las metodologías tradicionales (Ayuso, et al. 2021). Este documento examina los avances actuales centrándose en el área EGA (Figura 1).

El artículo analiza el potencial de la interacción entre la IA y el ser humano en el ámbito arquitectónico, enfocándose en sus aplicaciones en los procesos de diseño y las herramientas de representación. La sinergia entre la creatividad humana y la inteligencia de las máquinas abre nuevas posibilidades de colaboración, permitiendo a los profesionales abordar desafíos con una perspectiva renovada.

La primera parte del artículo se centra en la interacción entre dibujos realizados por un humano y modelos generativos de inteligencia artificial, con el objetivo de tener un control sobre los resultados obtenidos, reduciendo la aleatoriedad en los resultados. Se ha desarrollado un modelo de entrenamiento con Dreambooth para conseguir un estilo gráfico definido. En la segunda parte del artículo se han generado imágenes mediante modelos de difusión latente explorando seis estilos gráficos para tres ecosistemas arquitectónicos diferenciados, a los que hemos llamado biomas, emulando un proceso de diseño y dibujo en arquitectura: bocetos, dibujos con lápiz, acuarelas, dibujo técnico a tinta, modelado 3D y renderizado fotorrealista.

En el año 2022, Jason M. Allen ganó el primer premio en la categoría de arte digital en el Colorado State Fair con una obra titulada *Théâtre D'opéra Spatial* (Figura 2), una pintura generada completamente por IA utilizando MidJourney. Este concurso marcó uno de los primeros debates sobre la autenticidad del arte creado por inteligencia artificial y el papel del artista en este proceso.

El artista Jason M. Allen fue acusado de actuar de manera deshonesto y “hacer trampa”. En ese momento, no muchas personas habían utilizado estas herramientas y pocas conocían su funcionamiento. Los programas de texto a imagen como MidJourney o Stable Diffusion eran muy recientes y el popular ChatGPT aún no había sido lanzado. Desde entonces ha ocurrido un cambio notable: estas herramientas se han vuelto ampliamente populares y ha surgido una industria en torno a cómo utilizarlas de manera eficiente. En lugar de “hacer trampa”, esta habilidad ha recibido otro nombre: “prompt engineering” (ingeniería de indicaciones).

Introduction

Artificial intelligence (AI) has become a transformative force across various fields, and architecture is no exception. The integration of AI-powered human-machine interaction in architectural design and representation offers unprecedented opportunities to enhance creativity, efficiency, and precision. With the help of advanced algorithms and machine learning, architects and designers are now able to explore new and creative ways of designing, conceptualizing, and visualizing architecture.

As AI continues to evolve, its role in architecture is expected to grow, transforming the way we conceive and communicate ideas (Carpo 2023). From generating complex geometries to optimizing sustainable solutions, AI-driven systems enable architects to push the boundaries of traditional methodologies (Ayuso et al. 2021). This paper examines current developments with a focus on the area of Architectural Graphic Expression (EGA) (Figure 1).

The article analyzes the potential of human-AI interaction in the architectural field, focusing on its applications in design processes and representation tools. The synergy between human creativity and machine intelligence opens up new possibilities for collaboration, allowing professionals to tackle challenges from a renewed perspective.

The first part of the article focuses on the interaction between human-made drawings and generative artificial intelligence models, with the goal of gaining control over the outcomes and reducing randomness in the results. A training model using DreamBooth was developed to achieve a defined graphic style. In the second part, images were generated using latent diffusion models, exploring six graphic styles applied to three distinct architectural environments, referred to here as biomes, emulating an architectural design and drawing process: sketches, pencil drawings, watercolors, technical ink drawings, 3D modeling, and photorealistic rendering.

In 2022, Jason M. Allen won first prize in the digital art category at the Colorado State Fair with a piece titled *Théâtre D'opéra Spatial* (Figure 2), a painting entirely generated by AI using MidJourney. This competition marked one of the first debates about the authenticity of AI-generated art and the role of the artist in this process.



1



2



1. Imágenes generadas mediante un modelo de difusión latente (SDXL), basadas en dibujos a mano y optimizadas con un modelo de entrenamiento para controlar el estilo gráfico y color. Imágenes de los autores.
2. *Theatre D'Opera Spatial*. Jason M. Allen (Midjourney), 2022.
3. Generación de imágenes a partir de un mapa de segmentación. A) Nvidia GauGAN, 2019. B) Pix2Pix, 2016

1. Images generated using a latent diffusion model (SDXL), based on hand-drawn sketches and optimized with a training model to control graphic style and color. Images by the authors.
2. *Theatre D'Opera Spatial*. Jason M. Allen (Midjourney), 2022.
3. Image generation from a segmentation map. A) Nvidia GauGAN, 2019. B) Pix2Pix, 2016

Según avance el tiempo, saber cómo comunicarse con un programa de inteligencia artificial para obtener los mejores resultados será una habilidad básica, como enviar correos electrónicos o utilizar un motor de búsqueda en internet.

En los años recientes, aplicaciones como Chat GPT, Midjourney o DALL-E han popularizado el empleo de IA en diseño y arquitectura. Sin embargo, el empleo de IA generativa en arquitectura se remonta años atrás en el tiempo.

Evolución de los modelos de diseño generativo

La IA ha experimentado avances significativos en los últimos años, transformando la creación de contenido visual a través de modelos generativos y de difusión latente. Ian Goodfellow (2014) lanzó las Redes Generativas Antagónicas (GAN), que introdujeron un enfoque novedoso al enfrentar dos redes neuronales: un generador y un discriminador. Este marco permitió crear datos sintéticos que imitan ejemplos reales, sentando las bases para futuros avances en generación de contenido. En 2015, Google presentó DeepDream (Mordvintsev et al. 2015), una herramienta que amplificaba patrones visuales en imágenes, explorando cómo las redes neuronales procesan la información. Aunque inicialmente se diseñó para comprender mejor estas redes, DeepDream popularizó las imágenes surrealistas generadas por IA y abrió nuevas posibilidades en la creatividad asistida por tecnología.

En 2016, Philip Isola y su equipo introdujeron Pix2Pix (Isola, et al. 2016), un modelo que traduce imágenes de un dominio a otro utilizando redes neuronales, facilitando aplicaciones como la conversión de bocetos en imágenes realistas (Figura 3A). Un año después, en 2017, ampliaron esta idea con CycleGAN (Jun-Yan Zhu, et al. 2017), un sistema que permite la traducción de imágenes entre dominios sin necesidad de datos emparejados, lo que impulsó aplicaciones en la estilización artística y la adaptación de imágenes. En 2018, NVIDIA revolucionó los modelos generativos con StyleGAN, que ofreció un control granular sobre las características estilísticas de las imágenes generadas, permitiendo la creación de retratos hiperrealistas con un nivel de detalle sin precedentes. Posteriormente, en 2019, NVIDIA presentó GauGAN (Figura 3B), una herramienta innovadora que transformaba bocetos simples en paisajes realistas utilizando redes neuronales generativas, democratizando aún más el acceso a herramientas artísticas avanzadas.

Artist Jason M. Allen was accused of acting dishonestly and “cheating.” At the time, few people had used these tools, and even fewer understood how they worked. Text-to-image programs like Midjourney or Stable Diffusion were very new, and the now-popular ChatGPT had not yet been released. Since then, a notable shift has occurred: these tools have become widely popular, and an entire industry has emerged around how to use them effectively. What was once considered “cheating” has now acquired a new name: prompt engineering.

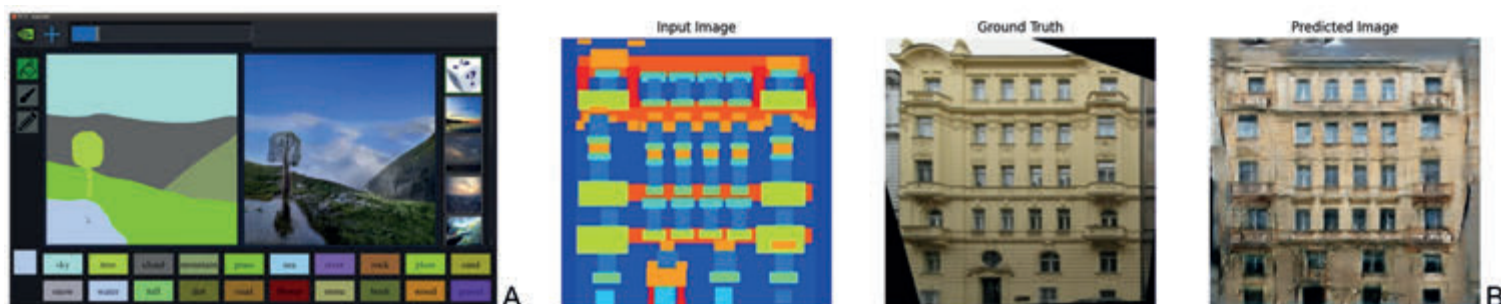
As time goes on, knowing how to communicate with an AI program to achieve optimal results will become a basic skill, just like sending emails or using an internet search engine.

In recent years, applications such as ChatGPT, Midjourney, and DALL-E have popularized the use of AI in design and architecture. However, the use of generative AI in architecture dates back several years.

Evolution of Generative Design Models

Artificial intelligence has undergone significant advances in recent years, transforming visual content creation through generative and latent diffusion models. In 2014, Ian Goodfellow introduced Generative Adversarial Networks (GANs), which brought a novel approach by pitting two neural networks against each other: a generator and a discriminator. This framework made it possible to create synthetic data that mimics real examples, laying the foundation for future breakthroughs in content generation.

In 2015, Google introduced DeepDream (Mordvintsev et al., 2015), a tool that amplified visual patterns in images while exploring how neural networks process information. Although





initially developed to better understand neural networks, DeepDream popularized surreal, AI-generated images and opened new possibilities for technology-assisted creativity. In 2016, Philip Isola and his team introduced Pix2Pix (Isola et al., 2016), a model that translates images from one domain to another using neural networks, enabling applications such as converting sketches into realistic images (Figure 3A). A year later, in 2017, they expanded this concept with CycleGAN (Jun-Yan Zhu et al., 2017), a system that allows image translation between domains without the need for paired data, boosting applications in artistic stylization and image adaptation.

In 2018, NVIDIA revolutionized generative models with StyleGAN, which offered granular control over the stylistic features of generated images, enabling the creation of hyper-realistic portraits with an unprecedented level of detail. Later, in 2019, NVIDIA introduced GauGAN (Figure 3B), an innovative tool that transformed simple sketches into realistic landscapes using generative neural networks, further democratizing access to advanced artistic tools.

DALL-E 1, launched in 2021 by OpenAI, was the first artificial intelligence model capable of generating images from textual descriptions using a transformer-based approach. Although DALL-E 1 was impressive in its ability to generate coherent and creative images, its quality and fidelity were limited, especially with complex descriptions. With the release of DALL-E 2 in 2022, OpenAI significantly improved image quality, enabling the generation of more realistic and detailed visual representations based on text prompts. DALL-E 2 also introduced new capabilities such as image editing (inpainting), allowing users to modify specific parts of an image while maintaining visual coherence. In 2023, DALL-E 3 took another leap forward, refining the interpretation of complex descriptions and further enhancing visual quality. This version enabled an even more precise and realistic image generation, as well as greater flexibility in editing.

In 2022, significant advancements were made in this field. MidJourney, developed by MidJourney Lab, stood out for its focus on creating stylized images from text and was widely adopted by communities of artists and designers.

With the release of MidJourney 2, also in 2022, improvements in image resolution and

DALL-E 1, lanzado en 2021 por OpenAI, fue el primer modelo de inteligencia artificial capaz de generar imágenes a partir de descripciones textuales, utilizando un enfoque basado en transformadores. Aunque DALL-E 1 impresionó por su capacidad para generar imágenes coherentes y creativas, su calidad y fidelidad eran limitadas, especialmente en descripciones complejas. Con la llegada de DALL-E 2 en 2022, OpenAI mejoró significativamente la calidad de las imágenes, permitiendo generar representaciones más realistas y detalladas de las descripciones de texto. DALL-E 2 también introdujo nuevas capacidades, como la edición de imágenes (inpainting), donde los usuarios podían modificar partes específicas de una imagen mientras mantenían la coherencia visual. En 2023, DALL-E 3 dio otro salto, perfeccionando la interpretación de descripciones complejas y mejorando aún más la calidad visual. Esta versión permitió una generación aún más precisa y realista, además de ofrecer una mayor flexibilidad en la modificación de imágenes.

En 2022 se produjeron avances significativos en este campo. MidJourney, desarrollado por MidJourney Lab, destacó por su enfoque en la creación de imágenes estilizadas a partir de texto, siendo ampliamente adoptado por comunidades de artistas y diseñadores.

Con la llegada de MidJourney 2, también en 2022, se mejoró la resolución de las imágenes y la capacidad de capturar detalles complejos, lo que permitió obtener resultados más precisos y estilísticamente variados. Sus versiones posteriores han aumentado la coherencia en las composiciones visuales y un mejor manejo de las perspectivas y proporciones, ofreciendo resultados más equilibrados y realistas, la fidelidad a las descripciones textuales y la capacidad de experimentar con estilos gráficos diversos.

En 2022 Stability AI lanzó Stable Diffusion, que democratizó aún más el acceso a las tecnologías de difusión latente, al ser un modelo de código abierto. Esta aproximación ha permitido una personalización avanzada y un mayor control creativo. Stable Diffusion XL (Podell, et al. 2023), lanzado en 2023, representa una versión mejorada de este modelo, con una mayor resolución y detalles más finos en las imágenes generadas (Li Z. et al. 2024). Herramientas como Automatic1111 y ComfyUI surgieron como interfaces accesibles para experimentar con Stable Diffusion, ofreciendo tanto a usuarios técnicos como creativos una plataforma para personalizar y explorar el flujo de trabajo de generación de imágenes.

En 2021 (Edward J. Hu, et al. 2021) se publicó LoRA (Low-Rank Adaptation), un método de optimización diseñado para adaptar modelos de inteligencia artificial de gran tamaño y pre-entrenados, como los modelos de difusión. Este enfoque permite ajustar los modelos utilizando una cantidad reducida de parámetros, lo que mejora significativamente su eficiencia. LoRA comenzó a ganar más popularidad y aplicación en 2022, debido a su capacidad para adaptar modelos como Stable Diffusion sin necesidad de reentrenar modelos completos, lo que reduce significativamente los recursos computacionales necesarios.

En 2022, Google introdujo DreamBooth (Ruiz, et al. 2022), un sistema diseñado para personalizar modelos generativos mediante el entrenamiento con datos específicos, facilitando la creación de contenido adaptado a estilos, entornos o características concretas.



Como referencia, también en 2022 OpenAI presentó ChatGPT y posteriormente en 2023 Google lanzó Gemini. Se trata de modelos de lenguaje natural con capacidad para generar texto coherente y conversaciones fluidas, subrayando la convergencia entre el procesamiento del lenguaje y las aplicaciones creativas de la IA. A medida que la tecnología sigue evolucionando, estos modelos representan un paso importante hacia una mayor integración de la inteligencia artificial en la vida cotidiana.

Durante este año, en 2024, OpenAI y Google presentaron SORA y VEO respectivamente, dando un paso adelante en la evolución de los modelos de difusión latente para la generación directa de videos a partir de descripciones de texto.

Los modelos de difusión latente, como MidJourney, DALL-E o Stable Diffusion, han superado las limitaciones iniciales de las redes generativas adversarias (GAN), al operar en un espacio latente comprimido. Esto les permite transformar ruido aleatorio en imágenes detalladas y personalizables, redefiniendo el panorama de la creatividad computacional. Estas tecnologías han democratizado el acceso a herramientas avanzadas, ampliando las posibilidades artísticas y técnicas, y fomentando una interacción más estrecha entre la inteligencia artificial y la creatividad humana. Desde las primeras GAN hasta los modelos actuales, la IA ha revolucionado no solo el ámbito creativo, sino también el técnico y ético, planteando nuevos desafíos en términos de propiedad intelectual, accesibilidad y originalidad, mientras transforma la producción de contenido visual en la era digital.

Metodología

Los dibujos de este artículo se han realizado usando una tableta Wacom Intuos Pro combinada con el lápiz ultrasensible Wacom Pro Pen 2, conectada a un ordenador con GIMP instalado, un software de código abierto (*GNU Image Manipulation Program*).

Las imágenes finales se han generado mediante Stable Diffusion, MidJourney, DALL-E, Veras y LookX. Todas son herramientas basadas en modelos de difusión latente, una técnica de inteligencia artificial que transforma descripciones textuales en imágenes visuales. Esta tecnología utiliza algoritmos de aprendizaje profundo (*deep learning*) para comprender e interpretar los matices del lenguaje, transformando palabras en imágenes complejas que pueden abarcar desde representaciones realistas hasta obras imaginativas. Se han empleado Automatic1111 y Comfy UI como interfaces de SD.

Para generar un estilo gráfico definido y controlado, se ha desarrollado un modelo de entrenamiento con Dreambooth en Stable Diffusion. Los trazados dibujados por un humano sirven para controlar y definir el resultado final de la imagen, siendo el input inicial de las IA junto con las descripciones de texto.

En la segunda parte del artículo, se han generado imágenes utilizando modelos de difusión latente, aplicando seis estilos gráficos distintos a tres biomas con atmósferas diferenciadas, cada uno representado con estilos arquitectónicos únicos y características diferenciadas.

the ability to capture intricate details allowed for more accurate and stylistically diverse results. Later versions further enhanced visual coherence, better handling of perspective and proportions, improved fidelity to textual prompts, and offered greater freedom to experiment with different graphic styles.

In 2022, Stability AI launched Stable Diffusion, further democratizing access to latent diffusion technologies by releasing it as open-source. This approach enabled advanced customization and greater creative control. Stable Diffusion XL (Podell et al., 2023), released in 2023, is an improved version of the model, offering higher resolution and finer detail in generated images (Li Z. et al., 2024). Tools like Automatic1111 and ComfyUI emerged as accessible interfaces for experimenting with Stable Diffusion, providing both technical and creative users with platforms to customize and explore image generation workflows.

In 2021, LoRA (Low-Rank Adaptation) was introduced by Edward J. Hu et al., as an optimization method designed to adapt large pre-trained AI models, such as diffusion models. This technique enables model fine-tuning using a reduced number of parameters, significantly improving efficiency. LoRA gained popularity and broader application in 2022 thanks to its ability to adapt models like Stable Diffusion without the need to retrain the entire model, thereby reducing computational resource requirements.

In 2022, Google introduced DreamBooth (Ruiz et al., 2022), a system designed to personalize generative models by training them with specific datasets, making it easier to create content tailored to particular styles, environments, or characteristics.

As a reference point, OpenAI also launched ChatGPT in 2022, and Google followed with Gemini in 2023. These are natural language models capable of generating coherent text and maintaining fluid conversations, highlighting the convergence between language processing and the creative applications of AI. As technology continues to evolve, these models represent an important step toward greater integration of artificial intelligence into everyday life.

In 2024, OpenAI and Google introduced SORA and VEO, respectively, marking a new stage in the evolution of latent diffusion models, enabling direct video generation from textual descriptions.



Latent diffusion models such as MidJourney, DALL-E, and Stable Diffusion have overcome the initial limitations of Generative Adversarial Networks (GANs) by operating in a compressed latent space. This allows them to transform random noise into detailed and customizable images, redefining the landscape of computational creativity. These technologies have democratized access to advanced tools, expanding both artistic and technical possibilities, and fostering a closer interaction between artificial intelligence and human creativity. From the earliest GANs to the latest models, AI has revolutionized not only the creative realm but also the technical and ethical domains, raising new challenges related to intellectual property, accessibility, and originality, while reshaping visual content production in the digital age.

Methodology

The drawings in this article were created using a Wacom Intuos Pro tablet combined with the ultra-sensitive Wacom Pro Pen 2, connected to a computer running GIMP, an open-source software (GNU Image Manipulation Program). The final images were generated using Stable Diffusion, MidJourney, DALL-E, Veras, and LookX. All of these are tools based on latent diffusion models, an artificial intelligence technique that transforms textual descriptions into visual images. This technology uses deep learning algorithms to understand and interpret the nuances of language, transforming words into complex images that can range from realistic representations to imaginative artworks. Automatic1111 and ComfyUI were used as interfaces for Stable Diffusion.

To generate a defined and controlled graphic style, a training model was developed using DreamBooth in Stable Diffusion. The hand-drawn sketches served to guide and define the final output, acting as the initial input for the AI models along with textual descriptions.

In the second part of the article, images were generated using latent diffusion models, applying six different graphic styles to three biomes with distinct atmospheres, each represented by unique architectural styles and differentiated characteristics.

Precisión y control de los resultados en la interacción Humano-IA

Los modelos de difusión latente como MidJourney o DALL-E permiten generar imágenes impactantes a partir de una descripción de texto. La interacción suele producirse mediante aproximaciones a base de prueba y error hasta conseguir los resultados deseados, sin poder definir los elementos con precisión.

Sin embargo, en el ámbito del diseño puede ser necesario tener un nivel de control más preciso, ya sea por tratarse de fases más avanzadas de un proyecto, o por querer plasmar una idea concreta en un dibujo.

La capacidad de comunicarse con una máquina a través del lenguaje natural ha supuesto un avance sin precedentes. Al integrar esta interacción con el dibujo a mano, se consigue una experiencia humano-máquina que se asemeja más a los métodos tradicionales empleados por arquitectos y artistas. Esto no solo reduce la aleatoriedad en los resultados, sino que también facilita la materialización de ideas concretas de manera más eficiente.

La figura 4 ilustra el flujo de trabajo empleado para generar las imágenes, integrando dibujos a mano con un modelo de difusión latente, específicamente Stable Diffusion.

La secuencia de dibujos en la figura 4A muestra la evolución de los bocetos realizados para generar la imagen final (Figura 4D). Los primeros trazos carecen de suficiente detalle, lo que dificulta una interpretación precisa por parte de la inteligencia artificial, dando lugar a resultados aleatorios. Al refinar el dibujo inicial y añadir más información, la IA comienza a captar la intención detrás del diseño, aunque aparecen artefactos en las líneas de los perímetros de los edificios y en los grafismos. La incorporación de una jerarquía mediante espesores de línea facilita que la IA distinga mejor las intenciones del dibujo. El resultado final se logró añadiendo sombreados que permiten separar los distintos planos de profundidad.

La segmentación de una imagen utilizando colores planos ha demostrado ser una técnica eficaz para controlar la generación de imágenes, como se observa en la Figura 3. Sin embargo, aunque este método puede ser adecuado al trabajar con modelos tridimensionales que pueden crear este mapa automáticamente, resulta menos práctico cuando se dibuja manualmente, ya que requeriría la preparación de imágenes específicas. En este contexto, el uso de grafismos y sombreados se presenta como una alternativa más intuitiva y natural.

Los bocetos iniciales proporcionan control sobre la silueta de la ciudad y los edificios (Figura 4B). Sin embargo, para definir con precisión el grafismo, se desarrolló un modelo de entrenamiento utilizando Dreambooth de Google, lo que permitió obtener resultados consistentes (Figuras 4C y 5).

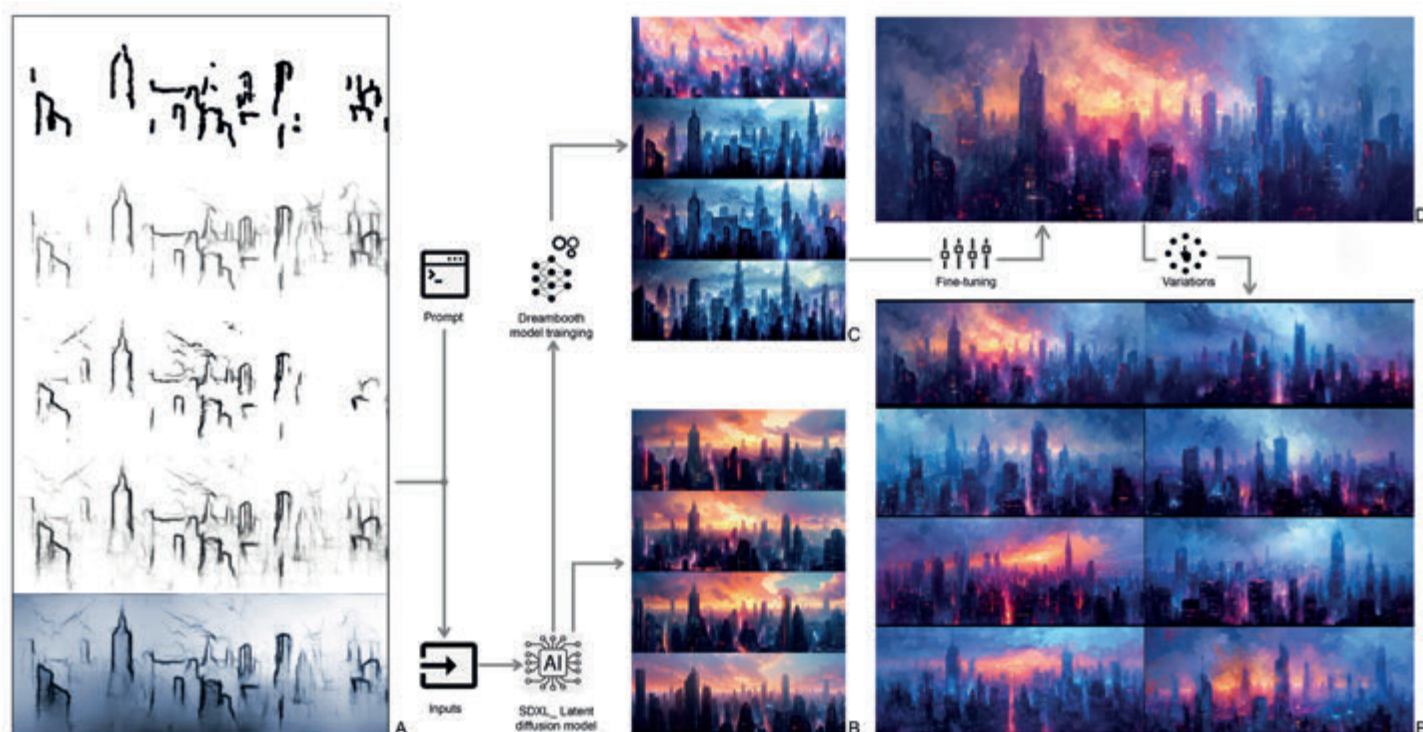
Todas las imágenes de este apartado se han generado usando la misma descripción de texto (prompt), resolución, ordenador y semilla del espacio latente, para reducir los parámetros que influyen en los resultados finales:

Descripción principal: *“Un paisaje urbano onírico al atardecer, con rasca-cielos de ventanas iluminadas, emergiendo de una ciudad vibrante con detalles texturizados, rodeados por una neblina mística. Paleta de colores fríos*



4. Imágenes generadas mediante un modelo de difusión latente (SDXL), basadas en dibujos a mano y optimizadas con un modelo de entrenamiento para controlar el estilo gráfico y color. Imágenes de los autores.

4. Images generated using a latent diffusion model (SDXL), based on hand-drawn sketches and optimized with a training model to control graphic style and color. Images by the authors.



4

dominada por azules y violetas, con suaves acentos de neón en magenta y naranja. Textura pictórica, que recuerda al arte impresionista, con pinceladas visibles y detalles abstractos. La perspectiva abarca una vista panorámica del horizonte urbano, con una sensación cinematográfica. En el estilo de federnetcity3.”

Especificaciones de la cámara: “Lente gran angular (capturar una vista panorámica). Baja profundidad de campo (efecto difuminado en los bordes). ISO 100 (reducir ruido en la imagen). Tiempo exposición = 2 segundos (acentuar las luces urbanas)”.

Descripción negativa: “Detalles borrosos, luces sobreexpuestas, caricatura, elementos fotorrealistas, nitidez excesiva, tonos sobresaturados, perspectiva distorsionada, texturas pixeladas.”

Trabajar únicamente mediante descripciones de texto puede generar resultados de gran calidad, pero con demasiada aleatoriedad y falta de control. Las imágenes de la figura 6 han utilizado la misma descripción de texto anterior, sin inputs adicionales.

La figura 7 muestra la superposición del boceto inicial (Figura 4A) con las distintas imágenes generadas a lo largo del proceso (Figuras 4B, 4C y 4D). Es posible observar la coincidencia en los perímetros de los edificios, aunque en algunas imágenes se han añadido estructuras adicionales que alteran el skyline de la ciudad. A lo largo del proceso, las líneas de grafismo han sido interpretadas de diversas maneras, transformándose en elementos como relámpagos, aves o nubes, hasta integrarse finalmente como parte de los trazos de la pintura en la composición final (Figura 5).

Accuracy and Control of Results in Human–AI Interaction

Latent diffusion models such as MidJourney or DALL-E allow for the generation of striking images from a simple text description. Interaction typically occurs through trial-and-error approaches until the desired result is achieved, without the ability to define elements with precision.

However, in the field of design, a higher level of control may be necessary, either because the project is at a more advanced stage, or because there is a need to express a specific idea through a drawing.

The ability to communicate with a machine using natural language represents an unprecedented advancement. When this interaction is combined with hand-drawn sketches, the human–machine experience more closely resembles traditional methods used by architects and artists. This not only reduces randomness in the results but also facilitates the realization of concrete ideas more efficiently.

Figure 4 illustrates the workflow used to



5



6



5. Imagen creada utilizando un modelo de difusión latente (SDXL), desarrollada a partir de un dibujo a mano y grafismos generados a partir de un modelo de entrenamiento. Imágenes de los autores.

6. Resultados aleatorios de imágenes generadas mediante un modelo de difusión latente, usando las mismas descripciones de texto. Imágenes de los autores.

5. Image created using a latent diffusion model (SDXL), developed from a hand-drawn sketch and graphic elements generated through a training model. Images by the authors.

6. Randomized image results generated using a latent diffusion model with identical text prompts. Images by the authors.



7. Coincidencia del dibujo inicial con las imágenes generadas. Imágenes de los autores.

7. Alignment of the initial sketch with the generated images. Images by the authors.



7

generate the images, integrating hand-drawn sketches with a latent diffusion model, specifically Stable Diffusion.

The sequence of drawings in Figure 4A shows the evolution of the sketches used to generate the final image (Figure 4D). The initial strokes lack sufficient detail, making it difficult for the artificial intelligence to interpret the design accurately, resulting in random outputs. As the initial drawing is refined and more information is added, the AI begins to grasp the design intent, although artifacts appear along the building outlines and in the graphic elements.

La superposición de las diferentes imágenes generadas en modo *overlay* (Figura 8) muestra el elevado nivel de control que este flujo de trabajo permite sobre las imágenes generadas. Se aprecia coincidencia en los volúmenes y perímetros de los edificios, y se mantiene la profundidad creada con los grafismos del dibujo inicial.

Aproximación a diferentes técnicas de representación gráfica mediante IA

Una gran ventaja de la generación de imágenes mediante IA es su capacidad de exploración, permitiendo la creación rápida de visualizaciones con una



8. Nivel de coincidencia; superposición de las imágenes generadas en modo *overlay*. De arriba abajo: dibujo inicial, superposición en escala de grises, superposición manteniendo los colores originales. Imágenes de los autores.

9. Imágenes generadas mediante un modelo de difusión latente. Exploraciones en la interconexión de las unidades habitables del bioma 3. Representaciones imitando dibujo a lápiz. Imágenes de los autores

8. Degree of alignment: overlay of the generated images. From top to bottom: initial sketch, grayscale overlay, overlay with original colors. Images by the authors.

9. Images generated using a latent diffusion model. Explorations of the interconnection between habitable units in Biome 3. Representations mimicking pencil drawings. Images by the authors.



8

Introducing a hierarchy through varying line thicknesses helps the AI better interpret the intentions of the drawing. The final result was achieved by adding shading, which separates the different depth planes.

Segmenting an image using flat colors has proven to be an effective technique for controlling image generation, as shown in Figure 3. However, while this method may be suitable when working with 3D models that can automatically generate such maps, it is less practical for hand-drawn sketches, as it would require the manual preparation of specific images. In this context, the use of graphic strokes and shading provides a more intuitive and natural alternative.

The initial sketches offer control over the silhouette of the city and the buildings (Figure 4B). However, to precisely define the graphic style, a training model was developed using Google's DreamBooth, which allowed consistent results (Figures 4C and 5).

All the images in this section were generated using the same text prompt, resolution, computer, and latent space seed, in order to minimize the parameters influencing the final results:

- Main description: "A dreamy urban landscape at sunset, with skyscrapers featuring illuminated windows, emerging from a vibrant city full of textured details, surrounded by a mystical haze. A cool color



9



palette dominated by blues and purples, with soft neon accents in magenta and orange. Painterly texture reminiscent of impressionist art, with visible brushstrokes and abstract details. The perspective offers a panoramic view of the urban skyline, with a cinematic feel. In the style of federnetcity3."

- Camera specifications: "Wide-angle lens (to capture a panoramic view). Low depth of field (blurred edges). ISO 100 (to reduce noise image). Exposure time = 2 seconds (to enhance urban lighting)."
- Negative prompt: "Blurry details, overexposed lights, cartoonish, photorealistic elements, excessive sharpness, oversaturated tones, distorted perspective, pixelated textures."

Working solely with text prompts can yield high-quality results, but often with excessive randomness and lack of control. The images in Figure 6 were generated using the exact same text prompt above, without any additional inputs.

Figure 7 shows the overlay of the initial sketch (Figure 4A) with the various images generated throughout the process (Figures 4B, 4C, and 4D). The outlines of the buildings align in most cases, although some images include additional structures that alter the city's skyline. Throughout the process, the graphic lines were interpreted in different ways, transforming into elements such as lightning, birds, or clouds, until they ultimately became part of the painterly strokes in the final composition (Figure 5).

The overlay of the different generated images (Figure 8) demonstrates the high level of control this workflow enables over the generated outputs. There is clear alignment in the volumes and outlines of the buildings, and the depth created by the graphic strokes in the initial sketch is preserved.

Approach to Different Graphic Representation Techniques Using AI

One major advantage of image generation through AI is its exploratory capacity, allowing for the rapid creation of visualizations in a wide range of graphic styles. The iterative nature of the method—largely based on trial and error—is particularly useful for reconstructing scenes from literary descriptions, visualizing nonexistent concepts, or developing architectural designs (Figure 9).

10. Imagen generada mediante un modelo de difusión latente. Representación del bioma 2, una ciudad con materiales autóctonos que se mimetizan en el desierto. Representación imitando una pintura al óleo. Imágenes de los autores.

10. Image generated using a latent diffusion model. Representation of Biome 2, a city using local materials that blends into the desert environment. Representation mimicking an oil painting. Images by the authors.

amplia gama de estilos gráficos. La naturaleza iterativa del método, basada en gran medida en prueba y error, es especialmente útil para la reconstrucción de escenas a partir de descripciones literarias, la representación de conceptos inexistentes o el diseño arquitectónico (figura 9).

Para la exploración y prueba en distintos estilos gráficos se han utilizado tres biomas diferenciados, acompañados de una descripción que define sus características: 1) Una ciudad futurista definida por su elevada densidad de población, que ha impulsado la construcción masiva de rascacielos. Un canal central atraviesa la ciudad, dividiéndola en dos partes. 2) Una ciudad ubicada en el desierto, destacada por sus estructuras abovedadas que emplean materiales locales, logrando una integración con el entorno (Figura 10). 3) Una ciudad submarina que sigue el estilo Art Nouveau, compuesta por una red de esferas traslúcidas ornamentadas con motivos inspirados en elementos naturales. Estas estructuras están interconectadas mediante pasarelas que enlazan los espacios habitables, reforzadas con arriostramientos que se anclan al fondo marino.

Para cada uno de los biomas se han generado seis estilos gráficos (croquis, dibujos a lápiz, acuarelas, dibujos técnicos a tinta, modelos tridimensionales y renderizados fotorrealistas), permitiendo una evolución gradual del diseño, incrementando el nivel de definición en cada etapa (figura 11). Cada una de estas variaciones se ha generado con diferentes sistemas de representación. La generación de una gran cantidad de imágenes durante este proceso sirve como base para nuevos proyectos creativos, facilitando la exploración y el ensayo en diversos estilos gráficos.

Aunque este enfoque puede sacrificar el control sobre los resultados debido a su componente aleatorio, esta característica también puede ser una ventaja al ofrecer inspiración y generar ideas novedosas en los diseños. La figura 12 muestra diferentes soluciones para las unidades habitables del bioma 3.

Como hemos visto, mediante el uso de IA podemos generar información sobre un boceto con estilos gráficos variados. Pero también podemos generar visualizaciones con distintos enfoques mediante modelos de difusión latente, explorando diversas alternativas que integren componentes aleatorios (figura 13).

Si quisiéramos lograr mayor consistencia en los resultados y una coincidencia en los dibujos con diferentes estilos gráficos, podríamos emplear las imágenes iniciales con menor nivel de definición como base de las nuevas representaciones, siguiendo el flujo mostrado en el apartado anterior. Alternativamente podríamos generar un mapa de segmentación de una imagen terminada, usándola como base de imágenes posteriores. Mayor control y precisión sobre los resultados implica mayor tiempo para su elaboración y gestión de los mapas necesarios.

A modo de resumen, la figura 14 presenta la combinación de imágenes generadas mediante diversas técnicas de representación gráfica aplicadas a los tres biomas, integrando distintos sistemas de representación.

Conclusiones

La inteligencia artificial ha alcanzado avances notables en la generación de



To explore and test various graphic styles, three distinct biomes were used, each accompanied by a description defining its characteristics: 1) A futuristic city characterized by high population density, which has driven the massive construction of skyscrapers. A central canal runs through the city, dividing it into two parts. 2) A desert city distinguished by domed structures made with local materials, achieving seamless integration with the environment (Figure 10). 3) An underwater city designed in the Art Nouveau style, composed of a network of translucent spheres decorated with motifs inspired by natural elements. These structures are interconnected by walkways linking the habitable spaces, reinforced with bracings anchored to the seafloor.

For each of the biomes, six different graphic styles were generated (sketches, pencil drawings, watercolors, technical ink drawings, 3D models, and photorealistic renderings) allowing for a gradual evolution of the design and an increasing level of detail at each stage (Figure 11). Each variation was created using different representation systems. The large number of images produced during this process serve as a foundation for new creative projects, facilitating exploration and experimentation across various graphic styles. Although this approach may sacrifice some control due to its random component, that very characteristic can also be an advantage, offering inspiration and generating novel design ideas. Figure 12 shows different solutions for the habitable units of Biome 3.

As we have seen, AI can be used to generate variations of a sketch in multiple graphic styles. But it can also create visualizations from different perspectives using latent diffusion models, exploring diverse alternatives that incorporate elements of randomness (Figure 13).

To achieve greater consistency across results and closer alignment between drawings in different graphic styles, one could use the initial low-definition images as the base for subsequent representations, following the workflow shown in the previous section. Alternatively, a segmentation map could be generated from a finished image and used as a foundation for further outputs. Higher control and precision require more time to prepare and manage the necessary segmentation maps.

As a summary, Figure 14 presents a combination of images generated through various graphic

representation techniques applied to the three biomes, integrating different systems of representation.

Conclusions

Artificial intelligence has made remarkable progress in image generation, enabling users to create striking visuals from simple text descriptions. However, control and precision over the results still present unresolved challenges, highlighting the need to develop new methods of interaction between humans and AI to improve consistency and accuracy.

These advancements have opened new possibilities for architects, designers, and artists, allowing them to express their ideas in innovative ways while saving time and resources. This capability not only democratizes access to creative tools but also fosters collaboration between human creativity and machine intelligence, ushering in a new era of artistic expression.

Latent diffusion models such as MidJourney, DALL-E, or Stable Diffusion represent an evolution beyond GANs. Instead of generating images directly, these models operate in a compressed latent space, using iterative processes to refine random noise into coherent visuals.

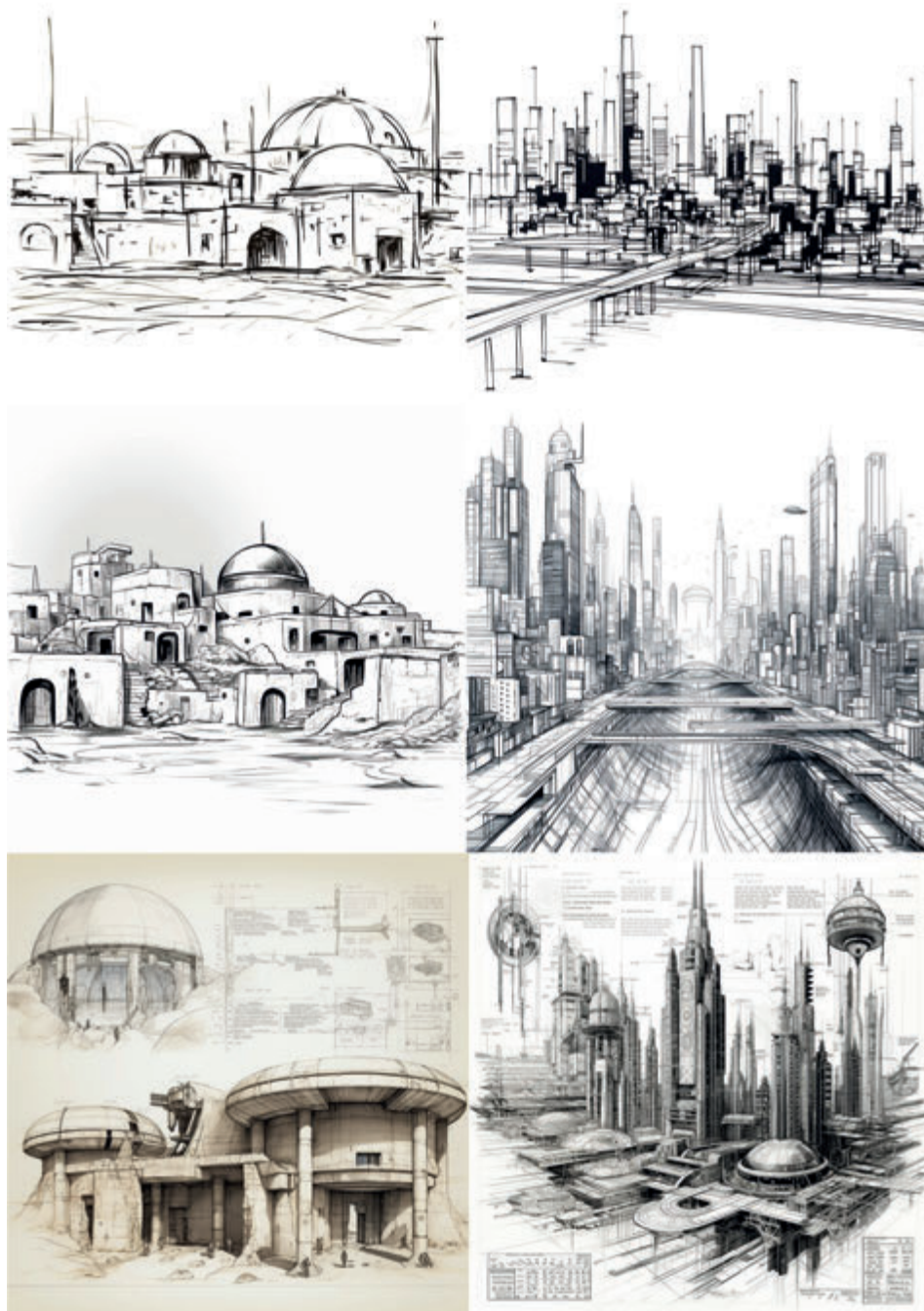
From Ian Goodfellow's GANs in 2014 to recent models like Stable Diffusion, AI has radically transformed how we understand and create visual content. The continuous evolution of these tools introduces new opportunities and challenges in terms of creativity, ethics, and accessibility.

The emergence of AI applications specifically tailored for architecture simplifies and accelerates workflows for users with less technical expertise. However, these tools still have limitations compared to more advanced latent diffusion models, especially when moving beyond photorealistic image generation or attempting to build customized workflows.

As these tools continue to evolve, they are likely to become even more sophisticated, enabling the generation of images that are not only visually compelling but also contextually relevant and tailored to users' specific preferences. The integration of AI into the creative process is reshaping our understanding of art and design, challenging traditional notions of authorship and originality

11. Imágenes generadas mediante modelos de difusión latente. Búsqueda de conceptos iniciales para los biomas 1 y 2. Representación imitando bocetos, con un aumento progresivo del nivel de detalle. Imágenes de los autores.

11. Images generated using latent diffusion models. Exploration of initial design concepts for Biomes 1 and 2. Representation mimicking sketch drawings, with a progressive increase in detail. Images by the authors.



11

imágenes, permitiendo a los usuarios crear visuales impactantes a partir de simples descripciones de texto. Sin embargo, el control y la precisión sobre los resultados aún representan desafíos por resolver, lo que resalta la necesidad de desarrollar nuevos métodos de interacción entre humanos e IA para mejorar el dominio y la exactitud en los resultados.

Estos avances han abierto nuevas posibilidades para arquitectos, diseñadores,

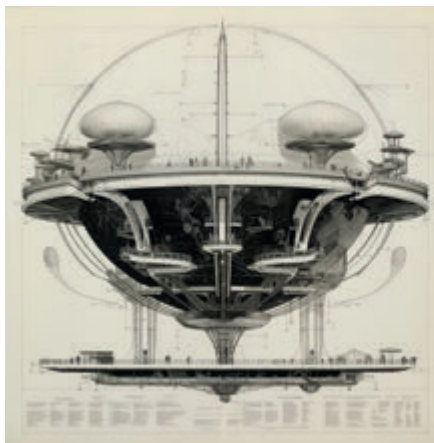


12. Imágenes generadas mediante modelos de difusión latente. Definición de diferentes unidades habitables del bioma 3. Representación imitando dibujos técnicos a tinta y render fotorrealista. Imágenes de los autores.

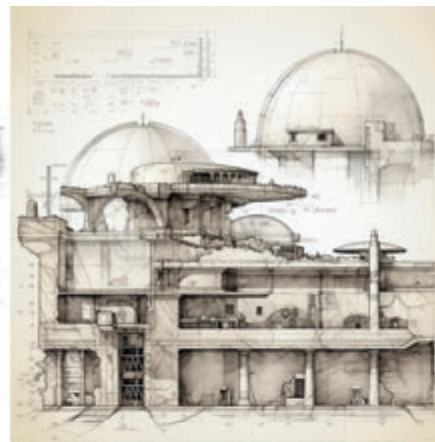
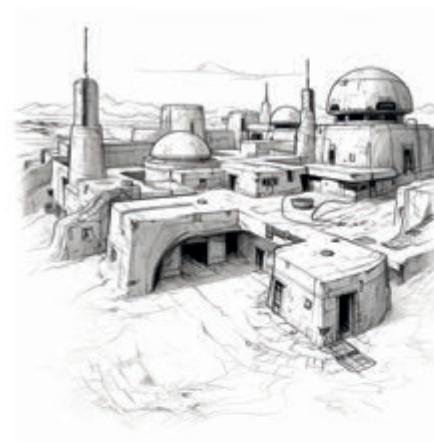
13. Imágenes generadas mediante modelos de difusión latente. Diferentes escenas de los biomas 2 y 3. Representación imitando dibujos a lápiz, acuarelas, dibujos técnicos a tinta y render fotorrealista. Imágenes de los autores.

12. Images generated using latent diffusion models. Definition of various habitable units in Biome 3. Representation mimicking technical ink drawings and photorealistic rendering. Images by the authors.

13. Images generated using latent diffusion models. Different scenes from Biomes 2 and 3. Representation mimicking pencil drawings, watercolors, technical ink drawings, and photorealistic rendering. Images by the authors.



12



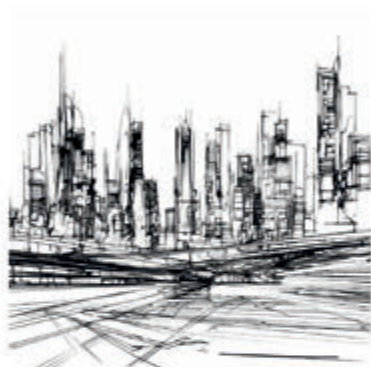
13

dores y artistas, permitiéndoles expresar sus ideas de formas innovadoras mientras ahorran tiempo y recursos. Esta capacidad no solo democratiza el acceso a herramientas creativas, sino que también fomenta la colaboración entre la creatividad humana y la inteligencia de las máquinas, dando lugar a una nueva era de expresión artística.

Los modelos de difusión latente, como MidJourney, DALL-E o Stable Di-

while inspiring new forms of storytelling and visual communication.

This transformation invites artists and designers to explore uncharted territories, where the boundaries between human creativity and machine-generated content begin to blur, encouraging new forms of

HUMAN
PERSPECTIVECONCEPT
SKETCH

PENCIL DRAWING



WATERCOLOR

AERIAL
PERSPECTIVE

FRONT VIEW



TOP VIEW



Skycraper city

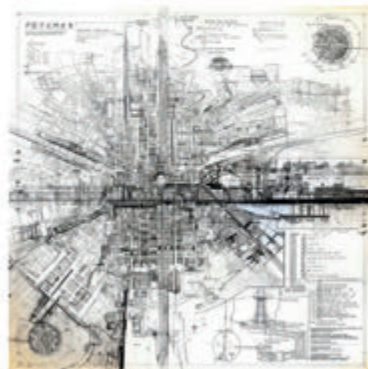
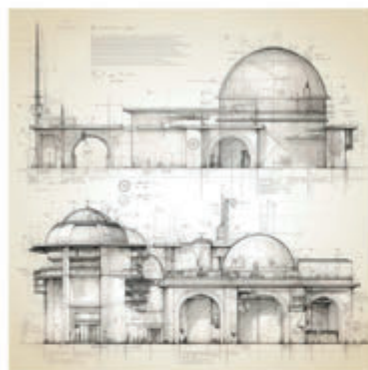
Deserted city



14. Tabla resumen de estilos gráficos. Representación de los tres biomas usando seis estilos gráficos con diferentes sistemas de representación. Imágenes de los autores.

14. Summary table of graphic styles. Representation of the three biomes using six graphic styles with different representation systems. Images by the authors.

TECHNICAL
INK DRAWING



3D MODEL



PHOTOREALISTIC
RENDER



Skyscraper city

Deserted city

Underwater city

Skyscraper city

Skyscraper city

Deserted city

Underwater city



collaboration that expands the boundaries of creativity.

This evolution not only broadens the set of tools available to creators but also raises important questions about the role of technology in art, prompting debates on copyright, ethics, and the future of creative expression. Who owns an image generated from a prompt? To what extent can a piece of content be considered original if it was produced by a model trained on millions of pre-existing images? Do these tools truly democratize creation, or do they introduce new forms of technological dependency? These are critical conversations that define the framework within which artists navigate this new landscape, ensuring that innovation is balanced with respect for intellectual property and individual creativity.

References

- Álvarez Ayuso, L., & del Blanco García, F. (2021). Aplicación de redes neuronales al diseño de vivienda colectiva. Procesos generativos de combinatoria y automatización mediante inteligencia artificial. *Rita_revista Indexada de Textos Académicos*, 16, 214–231. [https://doi.org/10.24192/2386-7027\(2021\)\(v16\)\(20\)](https://doi.org/10.24192/2386-7027(2021)(v16)(20))
- As, I., Pal, S. y Basu, P., 2018. Artificial intelligence in architecture: Generating conceptual design via deep learning. *International Journal of Architectural Computing*, 16(4), pp. 306–327. <http://doi.org/10.1177/1478077118800982>
- Aparicio Pérez, D., & del Blanco García, F. L. (2022). Design of immersive experiences for a utopian city. Graphic and virtual reconstruction of “The walking city”, Archigram. *EGA Revista de Expresión Gráfica Arquitectónica*, 27(44), 98–109. <https://doi.org/10.4995/ega.2022.14305>
- Bishop, C.M., 2006. *Pattern Recognition and Machine Learning*. Nueva York: Springer. ISBN 0387310738.
- García Ríos, P. y del Blanco García, F. L. (2023). New Babylon. Análisis y reconstrucción virtual de la visión utópica de Constant Nieuwenhuys, *EGA Expresión Gráfica Arquitectónica*, 28(47), pp. 256–271. <https://doi.org/10.4995/ega.2023.16173>
- Carpo, M., 2012. *The Digital Turn in Architecture 1992 - 2012*. Somerset: John Wiley & Sons
- Carpo, M., 2023. *Beyond Digital: Design and Automation at the End of Modernity*. Cambridge, MA: The MIT Press
- Carter, S., Armstrong, Z., Schubert, L., Johnson, I. y Olah, C., 2019. Activation Atlas. *Distill*, 4(3). DOI: <http://doi.org/10.23915/distill.00015>.
- Chaillou, S., 2019. *AI + Architecture, Towards a New Approach*. Harvard: Graduate School of Design.
- Del Blanco García, F.L.; González Cruz, A.J. (2024).

ffusion, representan una evolución respecto a las GAN. En lugar de generar imágenes directamente, estos modelos operan en un espacio latente comprimido, utilizando procesos iterativos para refinar ruido aleatorio hasta convertirlo en imágenes coherentes.

Desde las GAN de Ian Goodfellow en 2014 hasta los modelos más recientes como Stable Diffusion, la IA ha transformado radicalmente cómo entendemos y creamos contenido visual. La constante evolución de estas herramientas plantea nuevas oportunidades y desafíos en términos de creatividad, ética y accesibilidad.

La aparición de nuevas aplicaciones de IA específicas para arquitectura simplifica y acelera los flujos de trabajo para usuarios con menor especialización. Sin embargo, estas herramientas presentan limitaciones en comparación con los modelos de difusión latente más avanzados, especialmente si nos alejamos de la generación de imágenes fotorrealistas o intentamos desarrollar flujos de trabajo personalizados.

A medida que estas herramientas continúan evolucionando, es probable que se vuelvan aún más sofisticadas, posibilitando la generación de imágenes que no solo sean visualmente impactantes, sino también contextualmente relevantes y adaptadas a las preferencias específicas de los usuarios. La integración de la inteligencia artificial en el proceso creativo está transformando nuestra comprensión del arte y el diseño, desafiando las nociones tradicionales de autoría y originalidad, mientras inspira nuevas formas de narración y comunicación visual.

Esta transformación invita a artistas y diseñadores a explorar territorios inexplorados, donde los límites entre la creatividad humana y el contenido generado por máquinas se difuminan, fomentando una interacción que potencia el panorama creativo.

Esta evolución no solo amplía el conjunto de herramientas disponible para los creadores, sino que también plantea preguntas importantes sobre el papel de la tecnología en el arte, impulsando debates sobre derechos de autor, ética y el futuro de la expresión creativa. ¿A quién pertenece una imagen generada a partir de un *prompt*? ¿Hasta qué punto se puede considerar original un contenido producido por un modelo entrenado con millones de imágenes preexistentes? ¿Democratizan realmente la creación, o introducen nuevas formas de dependencia tecnológica? Estas conversaciones son cruciales, ya que configuran el marco dentro del cual los artistas navegan en este nuevo paisaje, asegurando que la innovación se equilibre con el respeto por la propiedad intelectual y la creatividad individual.

Referencias

- Álvarez Ayuso, L., & del Blanco García, F. (2021). Aplicación de redes neuronales al diseño de vivienda colectiva. Procesos generativos de combinatoria y automatización mediante inteligencia artificial. *Rita_revista Indexada de Textos Académicos*, 16, 214–231. [https://doi.org/10.24192/2386-7027\(2021\)\(v16\)\(20\)](https://doi.org/10.24192/2386-7027(2021)(v16)(20))



- As, I., Pal, S. y Basu, P., 2018. Artificial intelligence in architecture: Generating conceptual design via deep learning. *International Journal of Architectural Computing*, 16(4), pp. 306-327. <http://doi.org/10.1177/1478077118800982>
- Aparicio Pérez, D., & del Blanco García, F. L. (2022). Design of immersive experiences for a utopian city. Graphic and virtual reconstruction of “The walking city”, Archigram. *EGA Revista de Expresión Gráfica Arquitectónica*, 27(44), 98–109. <https://doi.org/10.4995/ega.2022.14305>
- Bishop, C.M., 2006. *Pattern Recognition and Machine Learning*. Nueva York: Springer. ISBN 0387310738.
- García Ríos, P. y del Blanco García, F. L. (2023). New Babylon. Análisis y reconstrucción virtual de la visión utópica de Constant Nieuwenhuys, *EGA Expresión Gráfica Arquitectónica*, 28(47), pp. 256–271. <https://doi.org/10.4995/ega.2023.16173>
- Carpo, M., 2012. *The Digital Turn in Architecture 1992 - 2012*. Somerset: John Wiley & Sons
- Carpo, M., 2023. *Beyond Digital: Design and Automation at the End of Modernity*. Cambridge, MA: The MIT Press
- Carter, S., Armstrong, Z., Schubert, L., Johnson, I. y Olah, C., 2019. Activation Atlas. *Distill*, 4(3). DOI: <http://doi.org/10.23915/distill.00015>.
- Chaillou, S., 2019. *AI + Architecture, Towards a New Approach*. Harvard: Graduate School of Design.
- Del Blanco García, F.L.; González Cruz, A.J. (2024). Development of Physical Simulations using Artificial Intelligence for Implementation in BIM Methodology. In: *Graphic Horizons*. EGA 2024. Springer Series in Design and Innovation, vol 43. Springer, Cham. https://doi.org/10.1007/978-3-031-57575-4_1
- Del Blanco García, F. L. (2022). Reconstructing Pérez Piñero' s Anoeta Velodrome. *Nexus Network Journal* 24, 913-934. <https://doi.org/10.1007/s00004-022-00590-3>
- Del Blanco García, F. L., Serrano Fernández, C. y González Cruz, A. J. (2024). Technology and Design in Cybernetic Architecture. Virtual Reconstruction of Cedric Price's Fun Palace. *EGA Expresión Gráfica Arquitectónica*, 29(51), pp. 242–257. doi: <https://doi.org/10.4995/ega.2024.20093>
- Goodfellow, I., et al., 2014. Generative Adversarial Nets. *Advances in Neural Information Processing Systems*, 27, pp. 2672-2680.
- Huang, W. y Zheng, H., 2018. Architectural Drawings Recognition and Generation through Machine Learning. *ACADIA 2018*, pp. 156-165.
- Kyle, S., Kat, P. y Adam, M., 2019. Fresh Eyes: A Framework for the Application of Machine Learning to Generative Architectural Design, and a Report of Activities at Smartgeometry 2018. *CAADRIA 2019*, Singapur, pp. 32-46.
- Mathias, M., et al., 2011. Automatic architectural style recognition. En: *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 171-176.
- Negroponte, N., 1970. *The Architecture Machine, Towards a More Human Environment*. Cambridge, Mass.: MIT Press. ISBN 0262640104.
- Negroponte, N., 1975. *Soft Architecture Machines*. Cambridge, Mass.: MIT Press. ISBN 9780262140188.
- Leach, N., 2021. *Architecture in the Age of Artificial Intelligence: An Introduction to AI for Architects*.
- Li, Z., et al. 2024. Hunyuan-DiT: A Powerful Multi-Resolution Diffusion Transformer with Fine-Grained Chinese Understanding. [arXiv:2405.08748v1](https://arxiv.org/abs/2405.08748v1)
- Luque-Sala, A., del Blanco García, F.L. (2023). A Virtual Reconstruction of Gaudi's Skyscraper Hotel Attraction Using Physics-Based Simulation. *Nexus Network Journal*, 25(3), 795-816. <https://doi.org/10.1007/s00004-023-00655-x>
- Ramesh, A., et al., 2022. Hierarchical Text-Conditional Image Generation with CLIP Latents. Disponible en: <https://arxiv.org/abs/2208.12242>.
- Podell, D., et al., 2023. SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis. *arXiv*, [cs.CV] 2307.01952v1. <https://doi.org/10.48550/arXiv.2307.01952>
- Zheng, H. y Huang, W., 2018. Understanding and Visualizing Generative Adversarial Networks in Architectural Drawings. *CAADRIA 2018*, Hong Kong.
- Development of Physical Simulations using Artificial Intelligence for Implementation in BIM Methodology. In: *Graphic Horizons*. EGA 2024. Springer Series in Design and Innovation, vol 43. Springer, Cham. https://doi.org/10.1007/978-3-031-57575-4_1