



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Arquitectura

La inteligencia artificial en la generación de imágenes de
arquitectura.

Aplicación directa de la Inteligencia Artificial en la
arquitectura y posterior aplicación de las herramientas de
modelado paramétrico para su ejecución material viable

Trabajo Fin de Grado

AUTOR/A: Eid-Masheh Casado, Yousef

Tutor/a: Molina Siles, Pedro Javier

CURSO ACADÉMICO: 2022/2023

Grado en Fundamentos de la Arquitectura

AUTOR/A: Eid-Masheh Casado, Yousef

Tutor/a: Molina Siles, Pedro Javier

CURSO ACADÉMICO: 2022/2023

An abstract architectural graphic on the right side of the page, consisting of numerous thin, curved, parallel lines that create a sense of depth and movement, resembling a modern building facade or a sculptural form. The lines are light gray and set against a white background.

La inteligencia artificial en la generación de imágenes de arquitectura

Aplicación directa de la Inteligencia Artificial en la arquitectura y posterior aplicación de las herramientas de modelado paramétrico para su ejecución material viable

Yousef Eid-Masheh Casado

Tutor: Pedro Molina-Siles
Dto. expresión gráfica arquitectónica ETSA, UPV

La inteligencia artificial en la generación de imágenes de arquitectura.

Aplicación directa de la Inteligencia Artificial en la arquitectura y posterior aplicación de las herramientas de modelado paramétrico para su ejecución material viable.

Los avances tecnológicos en la informática están cambiando la manera en la que los profesionales especializados en áreas creativas desarrollan sus proyectos. Las redes neuronales, conocidas comúnmente como inteligencia artificial, con su extensa base de datos, pueden realizar, con cierta precisión, tareas que solían estar reservadas para los expertos cualificados, convirtiéndolas en una cuestionada alternativa para el desarrollo creativo de estos profesionales.

En este trabajo de final de grado, reflexionaremos sobre los últimos descubrimientos en inteligencia artificial y aprendizaje automático aplicados a la arquitectura para poner en valor la trascendencia de estas tecnologías. Estudiaremos el proceso mediante el cual, partiendo de una imagen creada por un programa de inteligencia artificial (en la cual se aprecie una sinrazón constructiva, sin ejecución material viable), generaremos un modelo 3D haciendo uso de herramientas gráficas digitales de modelado paramétrico, obteniendo como resultado final una infografía 3D en la cual se aprecie una edificación racional.

Esta investigación, además, pretende demostrar que el uso de estas nuevas herramientas mejora significativamente el rendimiento y la capacidad de los profesionales, poniendo en evidencia la necesidad de adaptarse a los diversos cambios que se producirán en la industria en los próximos años.

Palabras clave: *Inteligencia Artificial, Arquitectura, Modelos de difusión, Render, Arquitectura paramétrica, Modelado paramétrico, Diseño paramétrico, Tecnología, Actualidad.*

Artificial intelligence for architecture images generation.

Straight application of Artificial intelligence models in architecture and subsequent implementation of parametric modelling tools to achieve a viable material execution.

The technological advances in computing are changing the way specialized professionals in creative areas develop their projects. Neural networks, commonly known as artificial intelligence, with its extensive database, can perform with certain precision tasks that used to be reserved for qualified experts, making them a questioned alternative for the creative development of these professionals.

In this final degree project, we will reflect on the latest discoveries in artificial intelligence and machine learning applied to architecture to highlight the transcendence of these technologies. We will study the process whereby, starting from an image created by an artificial intelligence model (in which a constructive irrationality is appreciated, without viable material execution), we will generate a 3D model using digital graphic parametric modelling tools, obtaining as a final result a 3D infography in which a rational edification is appreciated.

This research also aims to demonstrate that the use of these new tools significantly improves the performance and capacity of professionals, highlighting the need to adapt to the various changes that will take place in the industry in the coming years.

Key words: *Artificial intelligence, Architecture, Diffusion Models, Rendering, Parametric Architecture, Parametric Modelling, Parametric Design, Technology, Current affairs.*

La intel·ligència artificial en la generació d'imatges d'arquitectura.

Aplicació directa de la Intel·ligència Artificial en l'arquitectura i posterior aplicació de les eines de modelatge paramètric per a la seua execució material viable.

Els avanços tecnològics en la informàtica estan canviant la manera en la qual els professionals especialitzats en àrees creatives desenvolupen els seus projectes. Les xarxes neuronals, conegudes comúment com a intel·ligència artificial, amb la seua extensa base de dades, poden realitzar amb, una certa precisió, tasques que solien estar reservades per als experts qualificats, convertint-les en una qüestionada alternativa per al desenvolupament creatiu d'aquests professionals.

En aquest treball de final de grau, reflexionarem sobre els últims descobriments en intel·ligència artificial i aprenentatge automàtic aplicats a l'arquitectura per a posar en valor la transcendència d'aquestes tecnologies. Estudiarem el procés mitjançant el qual, partint d'una imatge creada per un programa d'intel·ligència artificial (en la qual s'aprecie una desraó constructiva, sense execució material viable), generarem un model 3D fent ús d'eines gràfiques digitals de modelatge paramètric, obtenint com a resultat final una infografia 3D en la qual s'aprecie una edificació racional.

Aquesta investigació, a més, pretén demostrar que l'ús d'aquestes noves eines millora significativament el rendiment i la capacitat dels professionals, posant en evidència la necessitat d'adaptar-se als diversos canvis que es produiran en la indústria en els pròxims anys.

Paraules clau: *Intel·ligència Artificial, Arquitectura, Models de difusió, Render, Arquitectura paramètrica, Modelatge paramètric, Disseny paramètric, Tecnologia, Actualitat.*

Índice

1. Introducción al proyecto	8
1.1 Introducción	9
1.2 Objetivos generales y específicos	10
1.3 Metodología y desarrollo	11
2. Estado del arte	12
2.1 Inteligencia Artificial	13
2.1.1 Inteligencia artificial en la arquitectura	14
2.1.2 Inteligencia artificial y sostenibilidad	16
2.1.3 Aprendizaje profundo	18
2.1.4 Modelos de generación de imagen	19
2.1.5 Ingeniería de prompts	22
2.2 Diseño paramétrico	24
2.3 Visualización 3D	27
3. IA para la generación de imágenes	28
3.1 Dall-e	30
3.1.1 Funcionamiento Dall-e	32
3.2 Midjourney	40
3.2.1 Funcionamiento Midjourney	42
3.3 Stable Diffusion	50
3.3.1 Funcionamiento Stable diffusion	52

4. Comparativa de resultados	62
4.1 Análisis comparativo	64
4.2 Conclusiones del análisis	69
5. Desarrollo del proyecto	70
5.1 Proceso de obtención de la muestra	72
5.2 Racionalización del objeto arquitectónico	76
5.2.1 Proceso de modelado	78
5.3 Resultado final	90
6. Conclusiones	92
Referencias bibliográficas	94
Bibliografía	96
Websites	97
Tabla de figuras	98
Anexo	100

1. Introducción al proyecto

1.1 Introducción

Las decisiones en cuanto a la forma de los edificios influyen en distintos aspectos, principalmente en lo relacionado a la arquitectura, la estética y la capacidad estructural de los mismos, todo ello ligado a la sostenibilidad. La forma influye en la iluminación y las pérdidas de carga térmica, pero también lo hace en el coste final y la superficie útil, por nombrar algunos ejemplos¹. Muchos de los trabajos de investigación en este campo actualmente se encuentran enfocados a la búsqueda de la generación de formas arquitectónicamente complejas tratando de optimizar al máximo estos parámetros.

Este trabajo, pretende poner de manifiesto los diversos avances tecnológicos que se están produciendo en los últimos años en el campo de la arquitectura, utilizados tanto como herramienta de soporte para los profesionales, como también de participante activo en los principales procesos de desarrollo, desde las fases creativas, asistiendo en la generación de referencias, hasta las fases de definición mediante la generación de planos e infografías, siendo capaces de realizar a lo largo del desarrollo del proyecto simulaciones de diversas índoles para la optimización de los recursos, consiguiendo así generar soluciones arquitectónicas más sostenibles.

Por otra parte, cada vez son más precisas las tecnologías empleadas para el análisis de los ecosistemas y de aspectos esenciales en la arquitectura, poniendo al alcance de los profesionales las herramientas necesarias para realizar análisis exhaustivos de las principales problemáticas de las ciudades, pudiendo abordarlas de manera localizada, fomentando así una mayor eficiencia de las mismas avanzando en el camino hacia la neutralidad climática de nuestros entornos.

1.2 Objetivos generales y específicos

Este trabajo tiene por objetivo estudiar y poner en evidencia la magnitud de los avances tecnológicos que se están produciendo en esta década y la necesidad de una adaptación de los profesionales a estos cambios, centrando la investigación en una de las múltiples áreas de impacto de la Inteligencia Artificial en el campo de la arquitectura: el proceso de diseño arquitectónico.

Dentro de este campo de gran amplitud, se desarrollará la aplicación de la IA en los primeros estadios de exploración formal del proyecto mediante la generación de referencias.

Por otro lado, se desarrollarán los siguientes objetivos específicos:

- Introducción a los conceptos básicos y principales aplicaciones de la Inteligencia artificial en la arquitectura.
- Comparación cualitativa y funcional de los principales programas de generación de imágenes del mercado.
- Desarrollo metodológico, que abarca desde el proceso de referenciación en una imagen producida mediante inteligencia artificial hasta la fase de modelado y concreción del objeto arquitectónico.

1.3 Metodología y desarrollo

El desarrollo de este trabajo de investigación se dividirá en dos fases: la primera parte de investigación se basa en una metodología teórica de recopilación de referencias bibliográficas mediante bases de datos y portales de búsqueda de artículos académicos para contextualizar los conceptos más influyentes de la investigación; para la segunda fase del proyecto, se parte de una metodología basada en la experimentación mostrando un proceso de comparación de recursos, exploración de las herramientas disponibles y selección y filtrado de las muestras obtenidas para la optimización de los resultados.

Para la búsqueda bibliográfica y desarrollo teórico del trabajo se emplearán los siguientes recursos:

- Polibuscador
- Google Academics
- EEE Xplore
- Research gate
- Arxiv.org

2. Estado del arte

2.1 Inteligencia Artificial

La Inteligencia Artificial (IA) es un campo de desarrollo que busca crear sistemas capaces de emular aspectos del comportamiento y pensamiento humano. Estos sistemas se esfuerzan por resolver problemas, tomar decisiones, comunicarse en lenguaje natural y adaptarse a nuevas circunstancias de manera inteligente, tratando de realizar tareas que, normalmente, requieren de la intervención humana .

Una de las características distintivas de la IA es su capacidad para aprender y mejorar su rendimiento a medida que interactúa con datos y experiencias, adaptándose a nuevas situaciones y entornos cambiantes. Al aprender de manera continua, los sistemas de IA pueden ajustar sus modelos y comportamientos para mantener un rendimiento óptimo incluso frente a desafíos imprevistos ².

Esta tecnología logra sus capacidades mediante el uso de algoritmos y modelos matemáticos complejos, especialmente a través del aprendizaje automático (machine learning). Esto permite que las máquinas detecten patrones y tendencias en grandes conjuntos de datos, mejorando sus predicciones y toma de decisiones con el tiempo. Además, la IA cuenta con el procesamiento del lenguaje natural (NLP), lo que le permite comprender el lenguaje humano, posibilitando interfaces intuitivas, asistentes virtuales conversacionales y un eficiente análisis de información textual.

La IA ha demostrado su potencial para revolucionar diversas áreas de la sociedad y la tecnología, con aplicaciones en atención médica, educación, industria y entretenimiento. Sin embargo, su rápida evolución también plantea desafíos éticos y sociales, como la privacidad de los datos y la automatización del trabajo. Es fundamental abordar estos aspectos de manera reflexiva y responsable para asegurar que la IA siga siendo una herramienta beneficiosa para la humanidad ³.

2.1.1 Inteligencia artificial en la arquitectura

La inteligencia artificial (IA) y el aprendizaje automático cada vez tienen más impacto en la industria de la arquitectura, “Buildtech” es el término acuñado para la referencia al nexo entre arquitectura y tecnología, actualmente se emplean soluciones tecnológicas para mejorar los procesos involucrados en el proyecto arquitectónico transformando la manera en que se conciben, crean y gestionan los edificios ⁴.

Entre las formas en las que la IA está siendo aplicada en el ámbito de la arquitectura cabe destacar ^{4,5}:

1. **Diseño Creativo:** La IA y el aprendizaje automático están permitiendo la creación de herramientas de diseño creativo, donde algoritmos evolutivos avanzados generan múltiples propuestas de diseño basadas en criterios específicos. Esto acelera el proceso de diseño y fomenta la exploración de soluciones innovadoras. Por otro lado, resulta posible el empleo de herramientas de generación de imágenes como referencia previa al proyecto para explorar nuevas posibilidades.
2. **Simulaciones Avanzadas:** Estas tecnologías impulsan simulaciones detalladas para comprender el comportamiento de edificios en términos estructurales, térmicos, acústicos y de flujo de aire. Estas simulaciones brindan una visión integral del rendimiento antes de la construcción, lo que facilita la toma de decisiones informadas y contribuye a la optimización de los recursos utilizados.

3. **Eficiencia Energética:** La IA y el aprendizaje automático analizan datos y patrones de uso para evaluar el consumo energético de edificios y sugieren mejoras para optimizar la eficiencia y reducir el gasto energético. Esto abarca desde la optimización de sistemas de iluminación hasta la gestión de climatización.
4. **Gestión de Proyectos:** Mediante algoritmos de aprendizaje automático, es posible analizar conjuntos extensos de datos relacionados con la planificación y ejecución de proyectos de construcción. Esto permite identificar patrones y optimizar la programación, el presupuesto y los recursos involucrados.
5. **Sensorización y Mantenimiento:** La inteligencia artificial es capaz de procesar datos de sensores en tiempo real para evaluar el estado y rendimiento de los edificios. Esto facilita el mantenimiento predictivo y la detección temprana de patologías, garantizando un funcionamiento eficiente y alargando la vida útil de los mismos.
6. **Diseño Contextualizado:** La IA y el aprendizaje automático analizan datos ambientales y de comportamiento de usuarios para influir en el diseño arquitectónico, adaptándolo al entorno y a las necesidades específicas de las personas.

2.1.2 Inteligencia artificial y sostenibilidad

La intersección entre la inteligencia artificial (IA) y la sostenibilidad ha dado paso a un panorama de posibilidades en la manera en que afrontamos los desafíos medioambientales y sociales de nuestro tiempo. La IA, con su capacidad para procesar y analizar enormes cantidades de datos, se ha convertido en una herramienta de gran alcance para promover prácticas sostenibles en diversos ámbitos, desde la gestión de recursos hasta la predicción de fenómenos naturales y la toma de decisiones estratégicas ⁶.

Un campo donde la IA está ejerciendo un impacto positivo es en la gestión de energía y recursos. Mediante algoritmos de IA, es posible optimizar la utilización de recursos energéticos y naturales en edificios y comunidades, identificando patrones de consumo y proponiendo estrategias para reducir el derroche y minimizar el impacto ambiental⁷. Esta optimización se traduce en una mayor eficiencia y en la disminución de la huella ecológica de nuestros edificios y ciudades.

Al analizar datos en tiempo real, como información satelital y datos climáticos, la IA puede anticipar eventos como inundaciones, incendios forestales y terremotos, jugando un papel crucial en la predicción y mitigación de desastres naturales⁸. Esto permite una respuesta más rápida y precisa ante situaciones críticas.

En el sector agrícola, la IA ofrece herramientas para monitorear y optimizar la producción de alimentos. Mediante la recopilación y análisis de datos, la IA puede prever cosechas, evaluar las condiciones del suelo y las necesidades de riego, y recomendar prácticas agrícolas más eficientes y amigables con el entorno. En la actualidad, podemos encontrar estudios que analizan en profundidad el flujo subterráneo de aguas y la contaminación de las mismas mediante aprendizaje automático⁹. Esto no solo aumenta la productividad, sino que también contribuye a minimizar la utilización de agroquímicos y a reducir los impactos negativos en los ecosistemas circundantes.

En las ciudades, la implementación de metodologías innovadoras de análisis y propuesta de soluciones respaldadas por inteligencia artificial cada vez son más habituales. En la actualidad, se están desarrollando proyectos como Plan Cero, en el cual se encuentran involucradas grandes entidades como la Universitat politècnica de Valencia y Arqueha arquitectura y urbanismo SL, entre otras. Este tipo de proyectos se encuentran enfocados a la mejora de aspectos que van desde la movilidad y la gestión de residuos, hasta la realización de simulaciones energéticas para tratar de localizar y mitigar las principales problemáticas de las ciudades acercándolas hacia el objetivo de la neutralidad climática.

2.1.3 Aprendizaje profundo

El deep learning, también conocido como aprendizaje profundo, es una subárea del campo más amplio del machine learning que se centra en la construcción y entrenamiento de modelos de inteligencia artificial que están basados en redes neuronales artificiales con múltiples capas. A través de estas capas, los modelos de deep learning pueden aprender automáticamente representaciones jerárquicas de los datos, lo que les permite realizar tareas de manera más eficiente y efectiva en comparación con otros enfoques tradicionales de machine learning.

El concepto fundamental detrás del deep learning es la idea de que las redes neuronales con múltiples capas pueden aprender automáticamente características y patrones complejos a partir de datos sin la necesidad de una extracción manual de características. Cada capa en una red neuronal aprende gradualmente características más abstractas y complejas a medida que se profundiza en la red, de ahí el término “profundo”¹⁰.

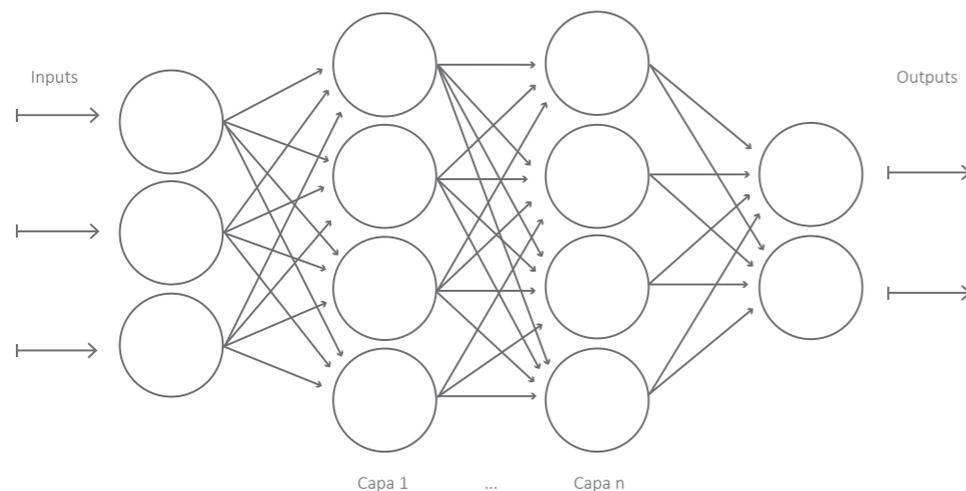


Fig 01. Redes neuronales

2.1.4 Modelos de generación de imágenes

Las Redes Neuronales Generativas Adversarias, comúnmente conocidas como GANs (por sus siglas en inglés, Generative Adversarial Networks) han recibido mucha atención en los últimos años. Las GANs fueron propuestas por primera vez por Ian Goodfellow y sus colegas en 2014¹¹ y, desde entonces, han demostrado ser una herramienta muy poderosa para diversas aplicaciones, entre ellas la generación de imágenes.

Como concepto fundamental, las GANs establecen una competencia entre dos redes neuronales: el generador y el discriminador.

El generador es una red neuronal profunda que toma como entrada un vector de ruido aleatorio y trata de generar datos que sean indistinguibles del conjunto de datos de entrenamiento que le han sido proporcionados. Inicialmente el generador produce datos aleatorios y sin sentido, pero a medida que avanzan los ciclos de entrenamiento, esta trata de generar datos que puedan engañar al discriminador para que los clasifique como reales.

El discriminador, por otro lado, es una red neuronal que trata de distinguir entre los datos reales del conjunto de entrenamiento y los datos producidos por el generador. Su tarea se basa en la asignación de una probabilidad alta a los datos reales y una probabilidad baja a los datos generados^{11, 12}.

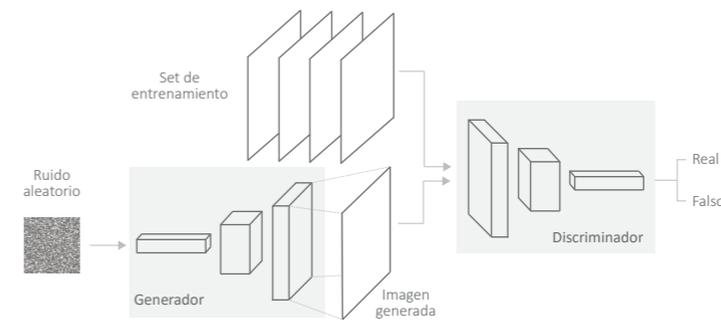


Fig 02. Redes generativas adversarias

En la actualidad han surgido nuevas técnicas, que por la gran calidad de los resultados que producen, han captado una atención significativa en el campo de la generación de imágenes. Entre ellas encontramos los modelos de difusión.

Los modelos de difusión representan una categoría novedosa de sistemas generativos que tienen la capacidad de generar imágenes de alta resolución. Su enfoque se basa en descomponer el proceso de generación de imágenes en una serie de pasos de reducción de ruido, lo que permite al modelo corregirse gradualmente, convergiendo en muestras de alta calidad y precisión. Cabe destacar que esta metodología iterativa resulta en velocidades de cálculo más lentas en comparación a otras técnicas generativas más tradicionales como las mencionadas con anterioridad ^{13,14}.

Estos modelos definen una cadena de pasos de difusión, donde cada paso depende únicamente del anterior, para añadir lentamente ruido aleatorio a los datos (Diffusion). Posteriormente, se procede a entrenar una red neuronal para que sea capaz de revertir este proceso a partir de un vector de ruido de entrada (Denoising), lo que habilita la generación de nuevos datos mediante la introducción de información condicional en dicho proceso ¹⁵.



Fig 03. Modelos de difusión

La generación de imágenes a partir de texto utilizando inteligencia artificial implica la creación de modelos que puedan comprender la descripción de texto y traducirla en una imagen coherente. Este proceso se basa en la combinación de técnicas de procesamiento del lenguaje natural (PLN) y redes neuronales generativas.

El primer paso es tomar la descripción en forma de texto como entrada, empleando técnicas de PLN para descomponer el texto en unidades significativas como palabras o frases y procesarlo, obteniendo como resultado una representación numérica del mismo con la que las redes neuronales puedan trabajar. Posteriormente, dicha codificación se introduce como condicionante para los modelos generadores con el fin de que esta influya en el resultado final ¹⁶.

Para que esto resulte efectivo, los modelos generadores han de estar previamente entrenados con una amplia base de datos formada por imágenes con texto asociado. Esto resulta altamente complicado y requiere una gran cantidad de recursos. Una de las compañías que ha conseguido poner solución a este problema ha sido Openai. Gracias a su tecnología CLIP (Contrastive Language-Image Pre-training)¹⁷, resulta posible asociar texto a extensos bancos de imágenes clasificándolas por categorías. Esta tecnología ha sido crucial para el entrenamiento de los modelos que estudiaremos posteriormente.



Fig 04. OpenAi Clip

2.1.5 Ingeniería de prompts

Autores como Sam Witteveen y Martin Andrews, han tratado de investigar acerca de como reaccionan los modelos de generación de imágenes ante la introducción de distintas combinaciones textuales. El concepto asociado a dicha práctica es ingeniería de prompts (Prompt engineering). La ingeniería de prompts se define como el proceso de diseñar y optimizar la entrada proporcionada a un modelo de procesamiento de lenguaje natural (PLN) para lograr los resultados deseados. Este proceso es particularmente importante al trabajar con modelos que generan datos basados en la entrada que reciben ¹⁸.

Un factor crítico en la ingeniería de prompts es la claridad y el contexto. Es fundamental que el prompt transmita de manera precisa la tarea o la información que se espera del modelo. Prompts ambiguos pueden llevar a respuestas confusas o incorrectas. Por lo tanto, la redacción cuidadosa del prompt es esencial para garantizar resultados relevantes, precisos y contextualmente apropiados ¹⁸.

Resulta de gran importancia tener en cuenta que el proceso de ingeniería de prompts implica iteración y experimentación. No siempre es posible obtener la respuesta deseada con el primer intento de prompt. Por lo tanto, ajustar y refinar los prompts según las respuestas obtenidas es un enfoque efectivo para lograr resultados óptimos.

Word placement and association

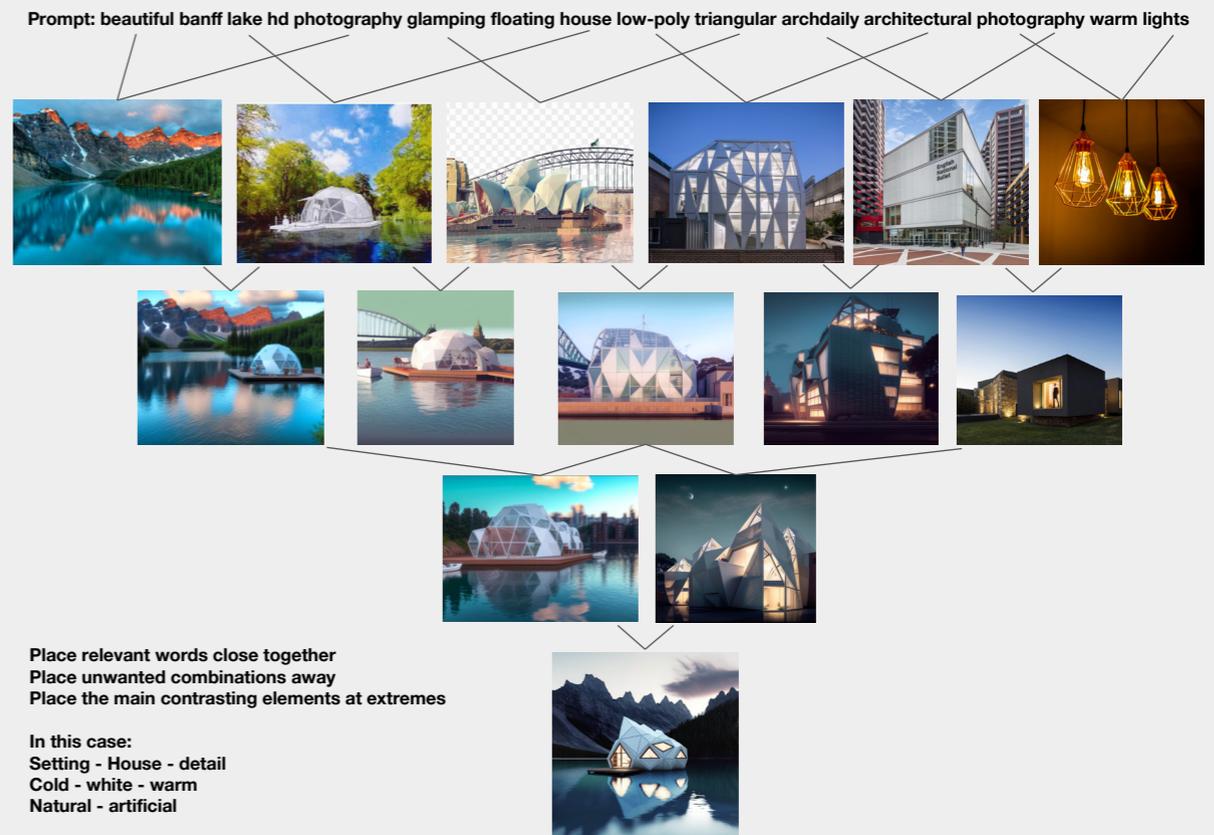


Fig 05. Ingeniería de prompts. MJ Studio 3

2.2 Diseño paramétrico

El diseño paramétrico en la arquitectura, representa una evolución significativa en los procesos de diseño en el campo de la arquitectura y planificación urbana. Este enfoque se basa en la idea fundamental de manipular parámetros y relaciones para generar formas y estructuras de manera dinámica y altamente adaptable.



Fig 06. Arquitectura paramétrica 1



Fig 07. Arquitectura paramétrica 2



Fig 08. Arquitectura paramétrica 3

El proceso de modelado paramétrico se materializa a través de componentes visuales que representan operaciones matemáticas y transformaciones geométricas. La clave reside en establecer conexiones paramétricas entre estos componentes, creando así un flujo de trabajo lógico y coherente. Al conectar la salida de un componente con la entrada de otro, se crea una red interconectada que permite que la información fluya y se ajuste en función de los cambios en los parámetros¹⁹.

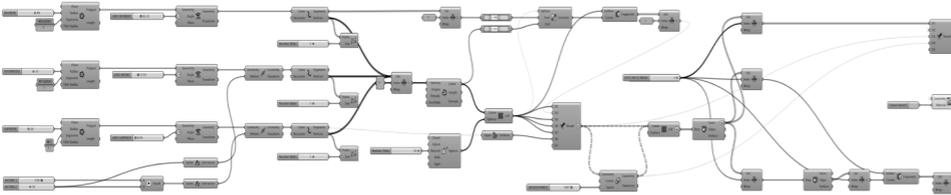


Fig 09. Algoritmo grasshopper

Una de las características distintivas del diseño paramétrico es la posibilidad de ajustar parámetros en tiempo real. Estos parámetros ajustables pueden representar dimensiones, ángulos, escalas y otros atributos de diseño. Al modificar uno de estos parámetros, el sistema responde instantáneamente, adaptando las formas y las relaciones según las instrucciones establecidas. Esto brinda a los arquitectos y diseñadores la capacidad de experimentar con múltiples configuraciones y explorar diversas alternativas de diseño en un entorno altamente iterativo.

Este enfoque también es altamente valioso para la optimización del diseño. Los usuarios pueden definir objetivos específicos y restricciones que guían el proceso de diseño. Al utilizar algoritmos, el sistema puede ajustar automáticamente los parámetros para lograr resultados que cumplan con los criterios predefinidos. Esto permite tomar decisiones más informadas y basadas en datos, optimizando las soluciones de diseño en función de consideraciones clave²⁰.

En resumen, el diseño paramétrico en la arquitectura, no solo simplifica y agiliza el proceso de diseño, sino que también amplía las posibilidades creativas. Al permitir la manipulación de parámetros y relaciones, los arquitectos pueden explorar, optimizar y generar formas y estructuras que respondan de manera dinámica a las necesidades y objetivos del diseño.

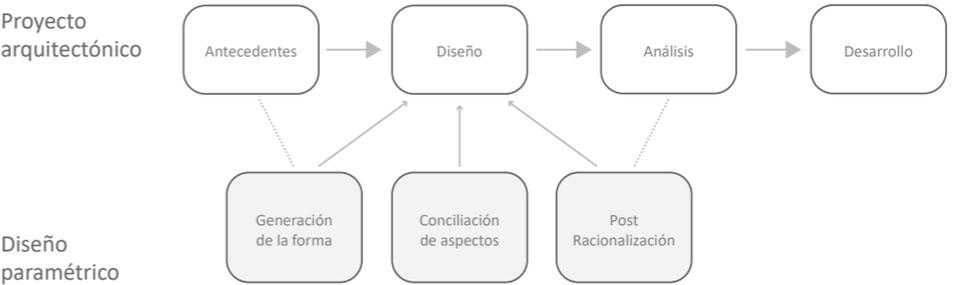


Fig 10. Diseño paramétrico

Para la realización de este trabajo, se emplearán herramientas de modelado paramétrico como Grasshopper 3D, un influyente plugin en el diseño paramétrico y arquitectura que nació como extensión para Rhinoceros 3D en 2008, desarrollado por David Rutten. Con el tiempo, ha evolucionado mediante múltiples versiones y actualizaciones, involucrando a diseñadores y arquitectos en su desarrollo y creando una red de intercambio de ideas. Grasshopper 3D se ha establecido como estándar en el diseño paramétrico, influyendo en la resolución de problemas en arquitectura y diseño.

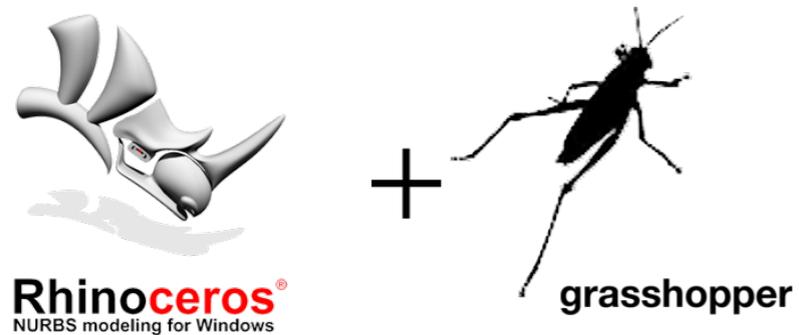


Fig 11. Programas de modelado paramétrico

2.3 Visualización 3D

La visualización arquitectónica a través de renderizado 3D es fundamental en el diseño y la arquitectura. Comienza con la creación de un modelo tridimensional usando software de modelado como 3ds Max, SketchUp o Rhinoceros 3D, donde los arquitectos y diseñadores dan forma y proporciones a las estructuras. Este modelo es la base esencial para la visualización final. Posteriormente, se avanza a la etapa de texturización, donde se añade realismo aplicando texturas a las superficies del modelo, que pueden ser imágenes fotográficas o generadas digitalmente, simulando diversos materiales.



Fig 12. Programas de visualización 3D

La iluminación juega un papel crucial en la creación de visualizaciones convincentes. Los programas de renderizado 3D permiten agregar fuentes de luz artificial y natural, ajustando intensidad, color y dirección para lograr efectos realistas, incluyendo iluminación global y reflejos que aportan profundidad a la escena. A medida que se avanza, se eligen cámaras virtuales y ángulos que resalten aspectos claves del diseño, similar a la selección de encuadres en fotografía, capturando vistas exteriores, interiores y detalles específicos.

El núcleo de la visualización arquitectónica es el proceso de renderizado, donde el software realiza cálculos complejos para simular interacciones de luz, generando efectos visuales como sombras, reflejos y transparencias, mediante algoritmos avanzados para alta fidelidad visual. Una vez finalizado el renderizado, las imágenes resultantes pueden llevarse a programas de posproducción, como Adobe Photoshop, para aplicar ajustes finales, como correcciones de color, efectos de desenfoque y atmosféricos, logrando la apariencia deseada. En conjunto, la visualización arquitectónica en 3D se convierte en una poderosa herramienta para comunicar, perfeccionar y presentar proyectos arquitectónicos de manera efectiva ²¹.

3. IA para la generación de imágenes

En la actualidad, existen diversas alternativas pioneras para la generación de imágenes mediante inteligencia artificial. Entre ellas destacan modelos que sobresalen por las tecnologías que emplean, resultando en una alta calidad en los resultados generados. Entre los modelos a analizar se encuentran Midjourney, DALL-E y Stable Diffusion. A continuación, se llevará a cabo un análisis comparativo de estas opciones con el propósito de determinar cual ofrece una mayor calidad y precisión a la hora de conseguir los resultados deseados. Por otro lado, se realizará un análisis comparativo de las funcionalidades que estos ofrecen a sus usuarios valorando, de la misma manera, la capacidad de aproximación a los resultados deseados.

3.1 Dall-e

Dall-e es un programa de generación de imágenes a partir de texto mediante inteligencia artificial, este es uno de los múltiples proyectos innovadores propuestos por la compañía Open AI, fundada en 2015 por Elon Musk y Sam Altman para investigar y desarrollar la inteligencia artificial ²². Esta compañía defiende que: *“Si se crea con éxito la inteligencia artificial generalizada (AGI), esta tecnología podría ayudarnos a elevar a la humanidad al aumentar la abundancia, dar un impulso extraordinario a la economía global, y contribuir al descubrimiento de nuevos conocimientos científicos que cambien los límites de la posibilidad.”* ²³.

Open AI es la responsable del desarrollo de múltiples modelos de inteligencia artificial junto con tecnologías que han permitido grandes avances en el sector, una de las tecnologías más trascendentes desarrolladas por esta empresa recibe el nombre de GPT (Generative pre-trained transformer). En base a este modelo de procesamiento del lenguaje natural, la empresa crea su tecnología CLIP (Contrastive Language–Image Pre-training), gracias a la cual resulta posible asociar texto a extensos bancos de imágenes ¹⁷.

El primer modelo de generación de imágenes entrenado con la tecnología CLIP y presentado junto a ella en 2021, fue Dall-e. Este modelo fue el primero con la capacidad de generar imágenes a partir de texto de forma fiel, llevando una de las facetas de la inteligencia artificial a manos de todo el mundo y consiguiendo la gran popularización de estos programas que encontramos en la actualidad. En septiembre de 2022, la compañía liberó la versión 2 de este programa, siendo esta la última actualización del programa en la actualidad ²².



Fig 13. Logo Dall-e
Fig 14. Resultado Dall-e, Pág.35



3.1.1 Funcionamiento

La plataforma para la generación de imágenes con Dall-e, se aloja en la página web oficial de Open Ai y, al emplear los recursos computacionales de la empresa para su funcionamiento, es un programa de pago. Su interfaz resulta intuitiva y fácil de aprender para el usuario, aunque por ello se encuentra muy limitada la personalización de las funciones del sistema. Cabe destacar que todas las imágenes generadas contienen una marca de agua en la esquina inferior derecha, el programa no ofrece la opción de retirarla, para ello sería necesario recurrir a programas de postproducción.

En primer lugar, es necesario introducir un texto para comenzar con la generación, El texto utilizado para la generación de este ejemplo será: Fotografía realista de una casa minimalista en el desierto.

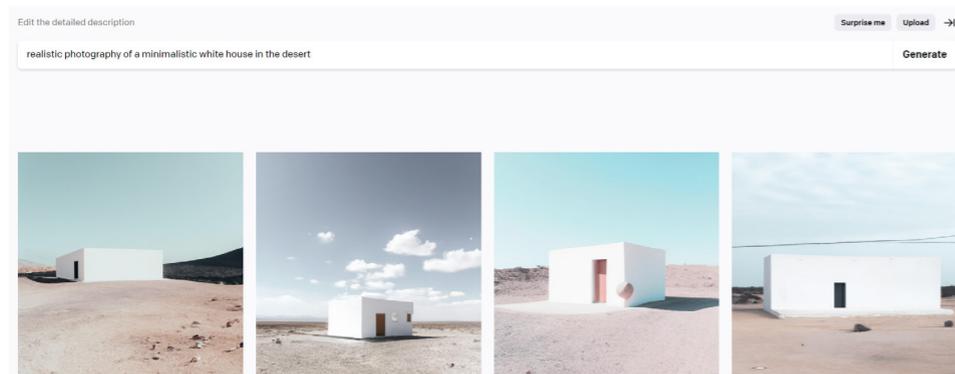


Fig 15. Funcionamiento Dall-e 1

El programa genera cuatro imágenes en base al texto introducido. Al ampliar cada una de ellas, el programa ofrece diversas opciones entre las cuales podemos encontrar: realizar variaciones, guardar y abrir en el editor la imagen deseada para realizar ajustes más precisos.

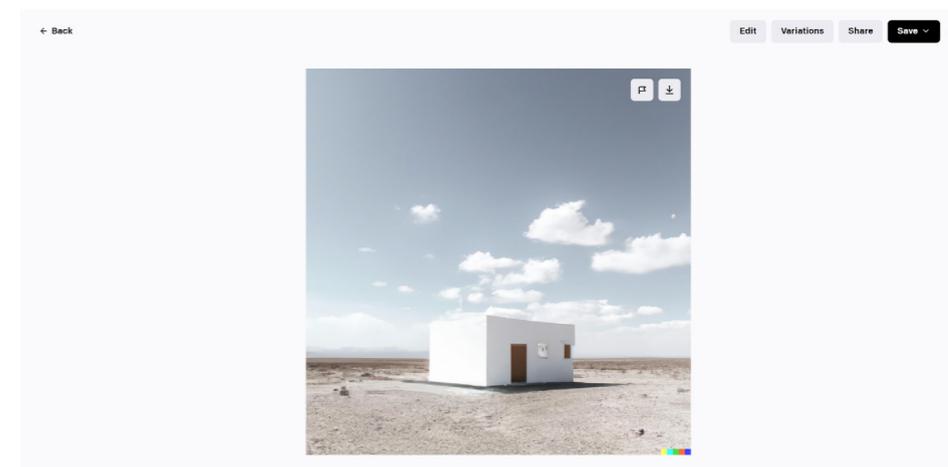


Fig 16. Funcionamiento Dall-e 2

Al solicitar variaciones de esta imagen el programa genera cuatro opciones más, este proceso se puede repetir indefinidamente hasta obtener un resultado que cumpla con las necesidades del usuario:

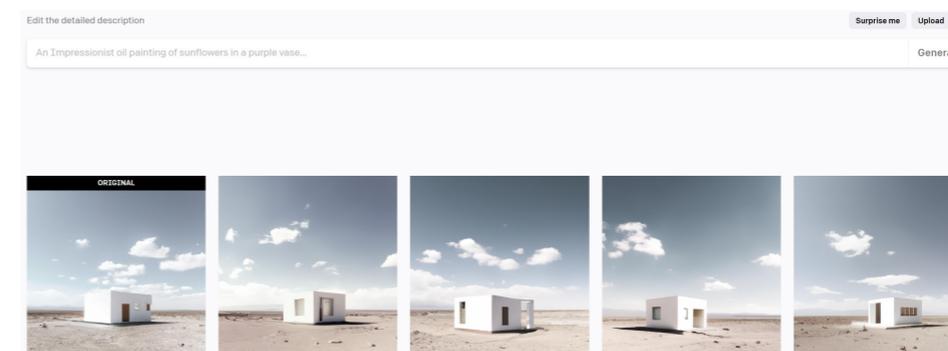


Fig 17. Funcionamiento Dall-e 3



Fig 18. Funcionamiento Dall-e 4

El apartado de edición, permite realizar tanto modificaciones de las imágenes generadas como de imágenes propias introducidas manualmente. Este programa tiene la particularidad de que permite, mediante la herramienta borrador, eliminar zonas concretas de la imagen sustituyéndolas por lo que se escriba en el cuadro de diálogo:

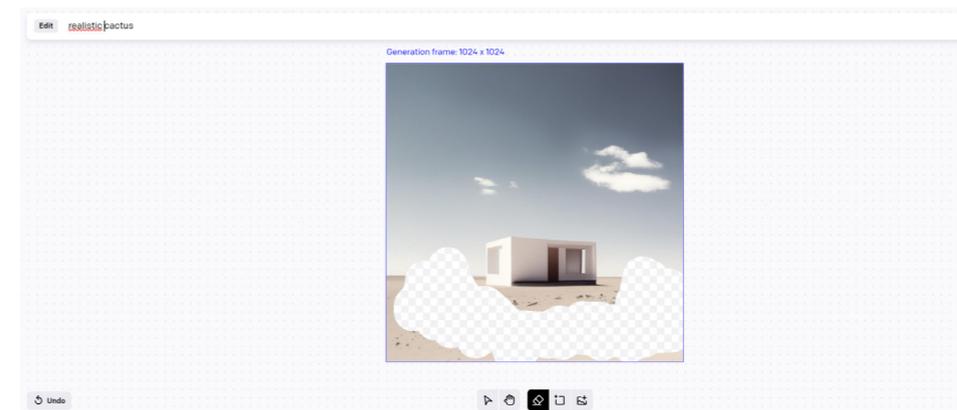


Fig 19. Funcionamiento Dall-e 5

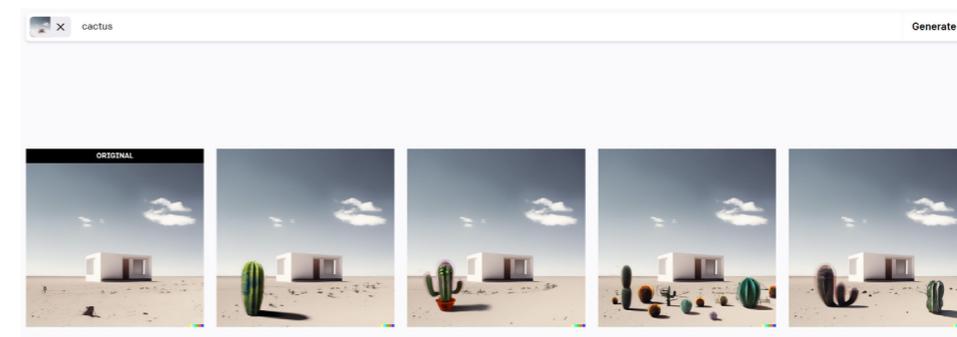


Fig 20. Funcionamiento Dall-e 6

Otra posibilidad que ofrece el programa es la de ampliar el marco de las fotografías admitiendo la introducción de texto para la generación de dicha ampliación:

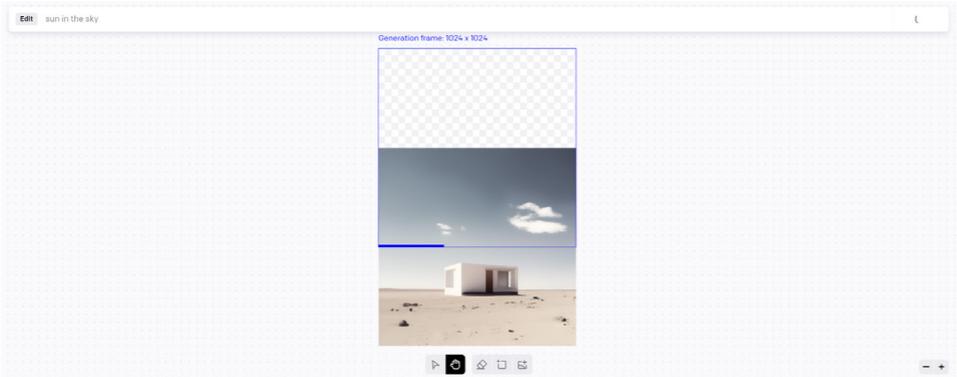


Fig 21. Funcionamiento Dall-e 7

Se generan diversas opciones, entre las que el programa permite la elección:

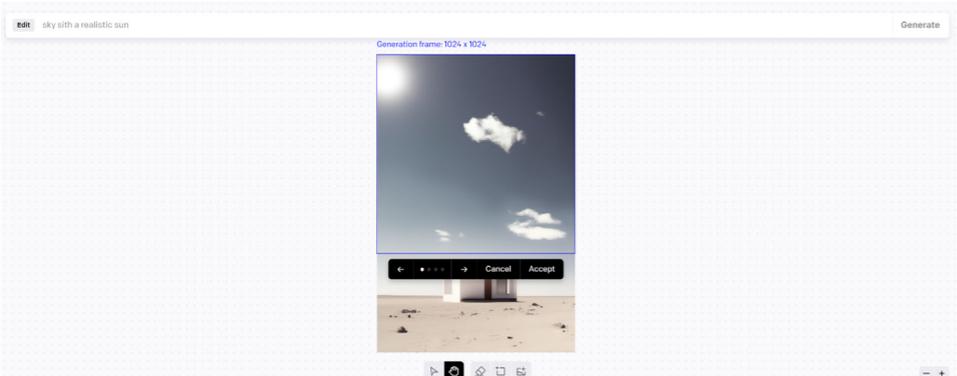


Fig 22. Funcionamiento Dall-e 8

La imagen de salida tiene una resolución final de 1024x1536 píxeles, esta resolución continúa siendo baja. Para ampliar esta imagen sin perder calidad, sería necesario utilizar otra inteligencia artificial que realice dicha tarea. En este caso se empleará un modelo de la compañía Nightmare AI.

Este modelo se ejecuta en la página web de Replicate, una plataforma que permite a los usuarios ejecutar programas de inteligencia artificial de código abierto. Para realizar el escalado de la imagen, se ha de seleccionar el archivo y el factor de ampliación de la imagen. En este caso se selecciona un factor de 2 buscando obtener una resolución de 2048x3072 píxeles.

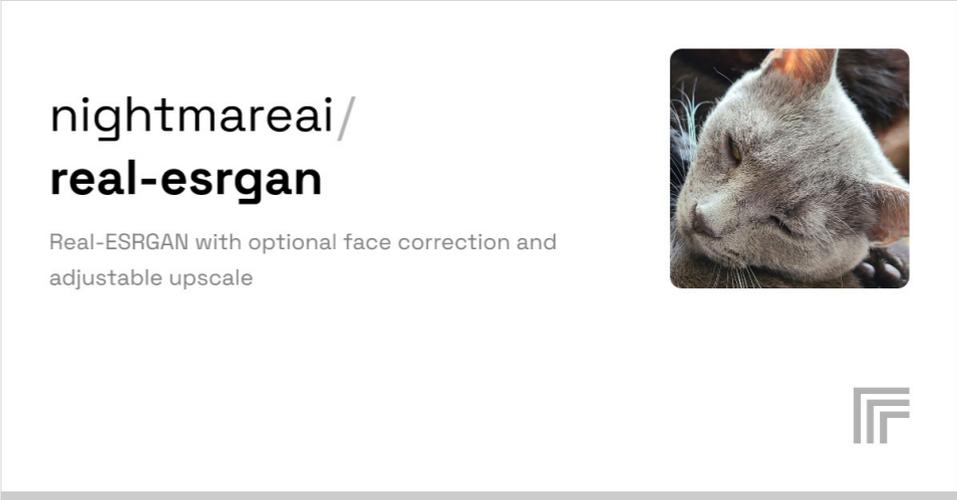


Fig 23. Programa de escalado

Los resultados obtenidos, siendo de mayor calidad, podrían ser ajustados mediante programas de edición de imagen, como Adobe Photoshop. Estos ajustes pueden incluir correcciones de color, ajustes de contraste y brillo, efectos de desenfoco para simular profundidad de campo y la adición de efectos atmosféricos para un mayor realismo.

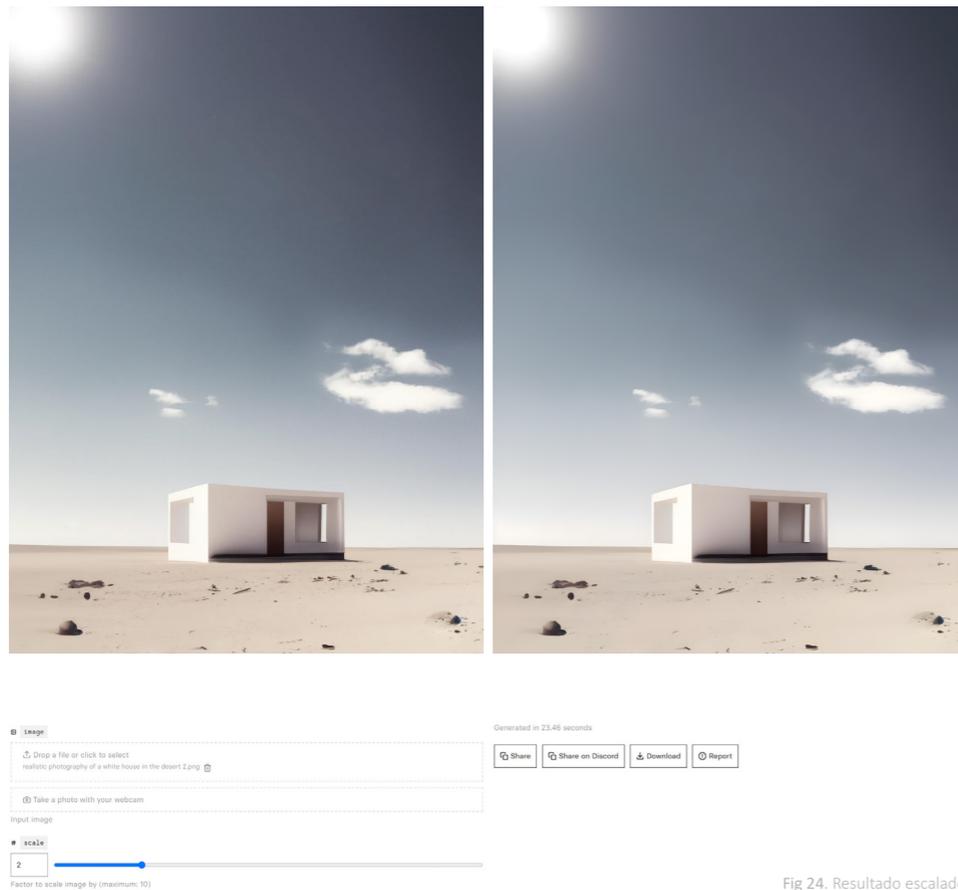


Fig 24. Resultado escalado

3.2 Midjourney

Midjourney es un laboratorio independiente de investigación que ha desarrollado un programa de inteligencia artificial que permite a sus usuarios generar imágenes a partir de descripciones textuales. El laboratorio, liderado por David Holz, es un pequeño equipo autofinanciado que se centra en el diseño, la infraestructura humana y la inteligencia artificial.

En Midjourney, los usuarios pueden crear imágenes a través de un bot alojado en Discord, una plataforma de comunicación en línea. Esta aplicación, entró en su primera fase de beta abierta el 12 de julio de 2022 y desde entonces la empresa ha seguido lanzando nuevas versiones. La versión 2 de su algoritmo se introdujo en abril de 2022, seguida por la versión 3 en julio del mismo año. Además, el 10 de noviembre de 2022, se presentó a los usuarios la iteración alfa de la versión 4, y el 15 de marzo de 2023 se lanzó la versión 5 en fase beta²⁴. Actualmente, el programa se encuentra en su versión 5.2.

El entrenamiento de Midjourney se basa en pares de imágenes y subtítulos extraídos de la base de datos LAION-5B. Este conjunto público, derivado de datos extraídos de Common Crawl en la web, comprende 5.000 millones de pares imagen-texto. Dichos datos fueron categorizados según el idioma, filtrados por resolución y evaluados en términos de presencia de marcas de agua y cualidades estéticas. LAION, una organización sin fines de lucro respaldada por Stability AI, fue la fuerza impulsora detrás de este conjunto de datos²⁵.



Fig 25. Logo Midjourney
Fig 26. Resultado Midjourney, Pág.45



3.2.1 Funcionamiento

La interfaz de generación de imágenes se encuentra bien estructurada, resultando intuitiva y fácil de usar para el usuario. Este programa se utiliza a través de un bot de discord, una aplicación de chat online muy conocida, este bot se encuentra completamente configurado limitando la capacidad de personalización, asegurando así la estabilidad del mismo a la hora de generar imágenes. De todas formas, resulta posible configurar ciertos parámetros mediante la introducción del comando `/settings` en el cuadro de diálogo.



Fig 27. Interfaz Midjourney

Estos ajustes incluyen parámetros como la selección del modelo a utilizar, la cantidad de estilización de la imagen, la cantidad de variación y la velocidad de generación de las mismas, entre otros. Mediante la introducción de comandos específicos, es posible seleccionar aspectos como la proporción de la imagen, la versión del programa y la semilla de ruido a utilizar.

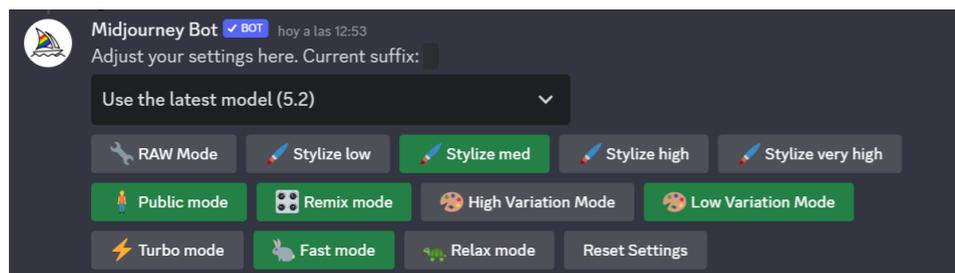


Fig 28. Interfaz Midjourney 2

Para introducir un prompt y comenzar la generación de imágenes, se teclea el comando `/imagine` en la conversación con el bot y a continuación el texto deseado. También resulta posible la introducción de imágenes que sirvan como referencia.



Fig 29. Interfaz Midjourney 2



Fig 30. Interfaz Midjourney 3

Una vez enviado, el programa procesa el texto y comienza la generación de cuatro imágenes, a medida que avanza la generación, se muestran los distintos pasos de reducción de ruido expresado en porcentaje:

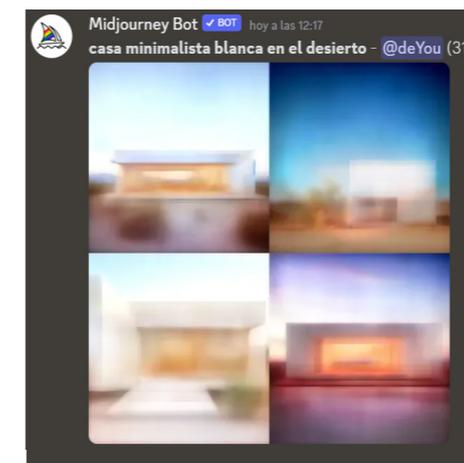


Fig 31. Interfaz Midjourney 4



Fig 32. Interfaz Midjourney 5

Tras completar la generación, el programa entrega cuatro imágenes de baja resolución ofreciendo diversas modificaciones a realizar, entre ellas encontramos variación y ampliado de cada una de las imágenes. Para conseguir mejores resultados, es conveniente realizar múltiples variaciones en los mismos pudiendo modificar el prompt para conseguir el resultado deseado.



Fig 33. Interfaz Midjourney 6



Fig 34. Interfaz Midjourney 7

Al realizar las variaciones de la primera imagen, se obtienen cuatro resultados similares a esta, pudiendo repetir este proceso indefinidamente hasta conseguir los resultados deseados. Ajustando el prompt en este momento, podrían realizarse variaciones más exageradas de la geometría, manteniendo aproximadamente la composición y la iluminación de la imagen seleccionada.

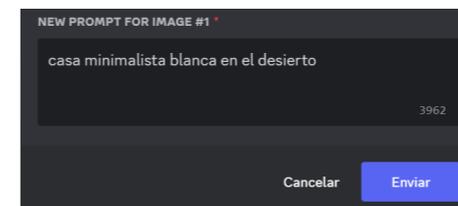


Fig 35. Interfaz Midjourney 8

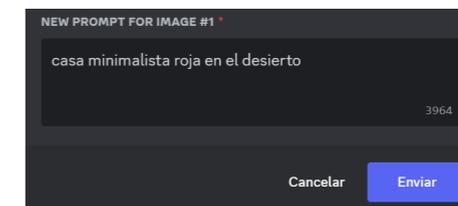


Fig 36. Interfaz Midjourney 9



Fig 37. Interfaz Midjourney 10



Fig 38. Interfaz Midjourney 11

Una vez conforme con el resultado obtenido, se procede a la ampliación de la imagen mediante la selección del botón de "Upscale" seguido del número de la imagen que se desea ampliar:

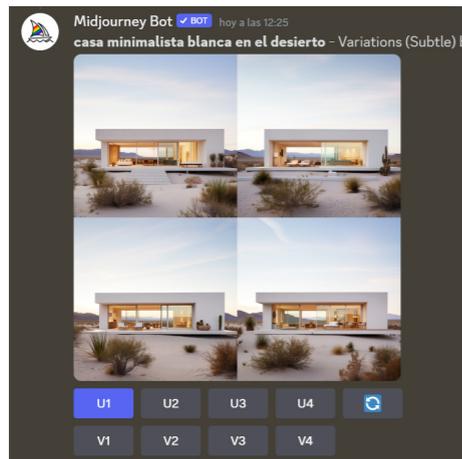


Fig 39. Interfaz Midjourney 12



Fig 40. Interfaz Midjourney 13

Como se puede observar, la imagen no se encuentra centrada, encontrando una parte del podio cortado. Esto puede solucionarse mediante la opción "zoom out":

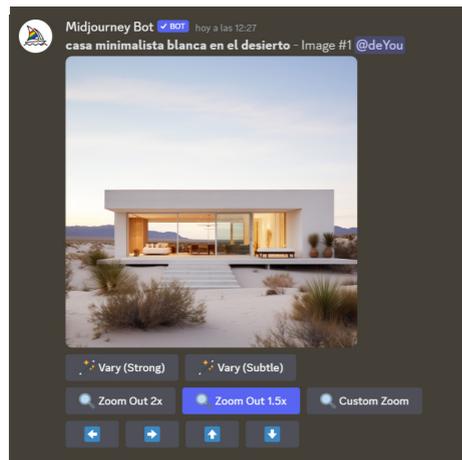


Fig 41. Interfaz Midjourney 14



Fig 42. Interfaz Midjourney 15

El programa permite, en este momento, modificar el formato de la imagen aumentando en un 50% uno de los lados de la imagen en la dirección seleccionada. Esta opción resulta de gran utilidad en los casos en los que no se realiza la selección previa de un formato distinto al predeterminado (1:1).

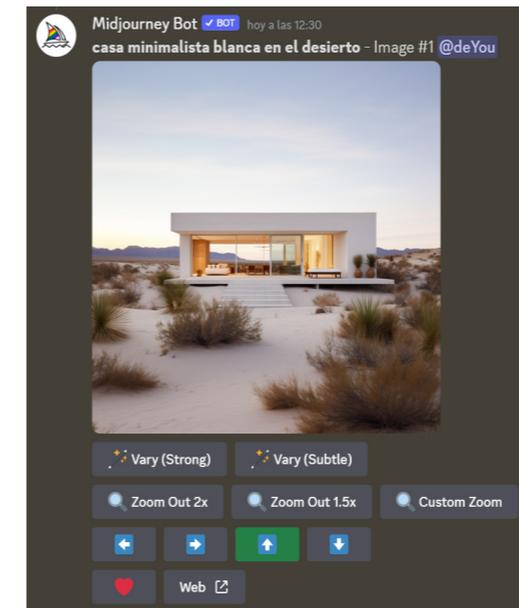


Fig 43. Interfaz Midjourney 16



Fig 44. Interfaz Midjourney 17

La imagen de salida tiene una resolución final de 1024x1536 píxeles, esta resolución continúa siendo baja. Para ampliar esta imagen sin perder calidad, sería necesario utilizar otra inteligencia artificial que realice dicha tarea. El modelo de escalado a utilizar será el desarrollado por la compañía Nightmare AI, empleado con anterioridad.

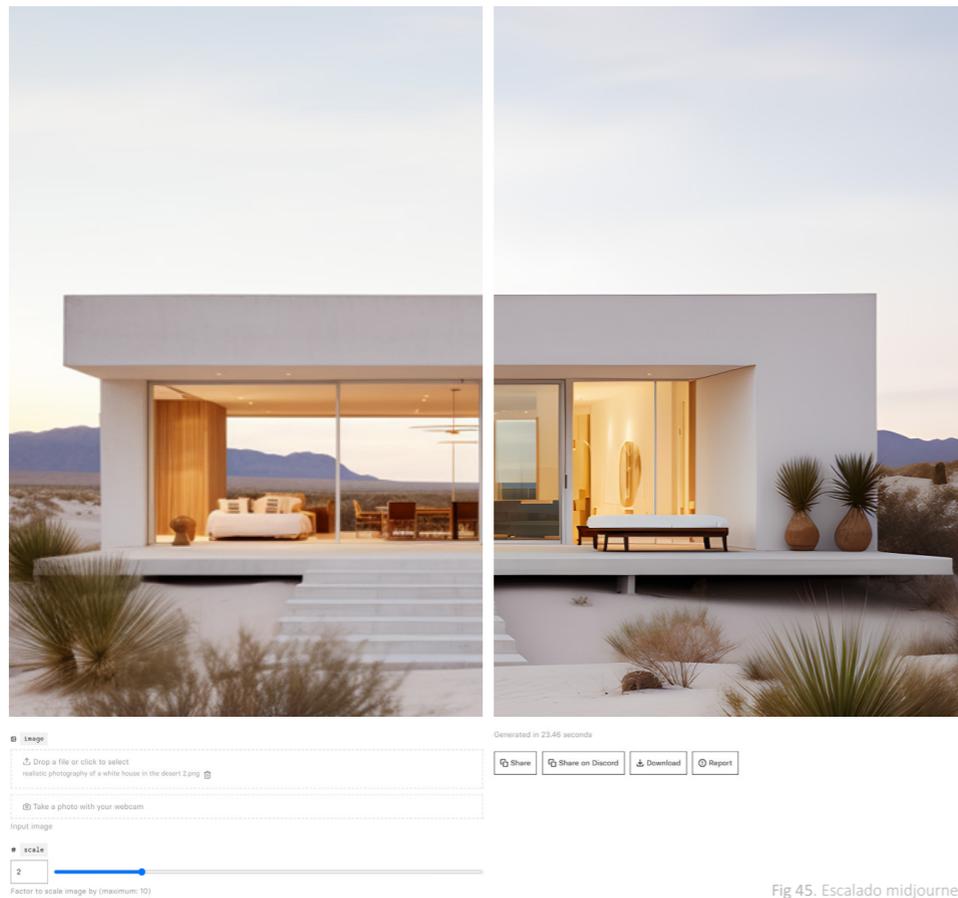


Fig 45. Escalado midjourney

Los resultados obtenidos, siendo de mayor calidad, podrían ser ajustados mediante programas de edición de imagen, como Adobe Photoshop. Estos ajustes pueden incluir correcciones de color, ajustes de contraste y brillo, efectos de desenfoco para simular profundidad de campo y la adición de efectos atmosféricos para un mayor realismo.

3.3 Stable Diffusion

Stable difusión surge como un modelo de aprendizaje automático, fruto de la colaboración entre Runway y LMU Múnich. Este modelo se utiliza para generar imágenes digitales de alta calidad a partir de descripciones de texto. Este tiene la particularidad de que, a diferencia de sus principales competidores, es gratuito. Stable Diffusion puede ser ejecutado en hardware convencional equipado con tarjetas gráficas (GPUs)²⁶.

El entrenamiento de Midjourney se basa en pares de imágenes y subtítulos extraídos de la base de datos LAION-5B. Este conjunto público, derivado de datos extraídos de Common Crawl en la web, comprende 5.000 millones de pares imagen-texto. Dichos datos fueron categorizados según el idioma, filtrados por resolución y evaluados en términos de presencia de marcas de agua y cualidades estéticas. LAION, una organización sin fines de lucro respaldada por Stability AI, fue la fuerza impulsora detrás de este conjunto de datos²⁷.

Las primeras versiones del modelo se desarrollaron utilizando imágenes de resolución 512 × 512. Sin embargo, se observa que la calidad de las imágenes generadas disminuye cuando la imagen demandada por el usuario es distinta a la “esperada” de 512 × 512. La versión 1.5, lanzada el 20 de octubre de 2022, presentó mejoras menores. Una significativa evolución se dio con la liberación de la versión 2.0 el 24 de noviembre de 2022, que permitió generar imágenes nativas con una resolución de 768×768. En abril de 2023, la compañía lanza la versión beta del modelo reentrenado por completo. Esta nueva versión se denomina Stable Duffusion XL ya que cuenta con un conjunto de entrenamiento mucho mayor y resulta capaz de generar imágenes con la resolución de 1 megapíxel (1024x1024)²⁸. En julio de 2023, se liberó la versión 1.0 de SDXL, última versión disponible en este momento y con la que se realizarán las pruebas posteriores.



Fig 46. Logo Stable Diffusion
Fig 47. Resultado Stable Diffusion, Pág.54



3.2.1 Funcionamiento

Al ser de código abierto, existe la posibilidad de ejecutar Stable diffusion en plataformas web como Playground o Leonardo Ai, pero estas plataformas son de pago ya que la generación de las imágenes consume recursos de sus servidores y no del ordenador del usuario. Para ejecutar este programa en un ordenador personal, hay varias opciones con distinto nivel de personalización.

Una de las formas más personalizables y óptimas de ejecutarlo es mediante un programa llamado Comfy UI. En dicha plataforma es posible cargar los modelos, previamente instalados en el ordenador, permitiendo generar imágenes de manera gratuita utilizando los recursos del propio equipo, siempre y cuando este cumpla con los requerimientos mínimos en cuanto a tarjeta gráfica (GPU).

Comfy UI es una plataforma basada en nodos que permite trabajar con distintos modelos de generación de imágenes permitiendo configuraciones completamente personalizadas, por lo que requiere un previo conocimiento del sistema. Esta capacidad de personalización, tiene como ventaja la posibilidad de ejecutar automáticamente todo el programa de una sola vez.

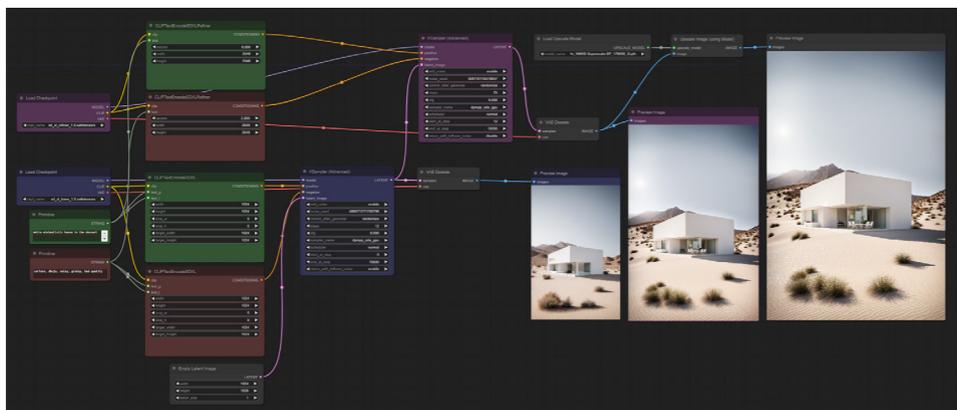


Fig 48. Funcionamiento Stable Diffusion

Antes de comenzar con la generación, es necesario introducir algunos parámetros para configurar el sistema. Estos parámetros se introducirán en los nodos correspondientes, siendo posible programar nodos propios o descargar nuevas funcionalidades creadas por la comunidad.

El primer paso será cargar los modelos de generación de imágenes, previamente descargados en el ordenador. Stable diffusion XL, se compone de un modelo base que genera una imagen de baja resolución y un modelo refinador que, partiendo del resultado extraído del anterior modelo, añade detalles y mejora la iluminación manteniendo la estructura principal de la imagen.

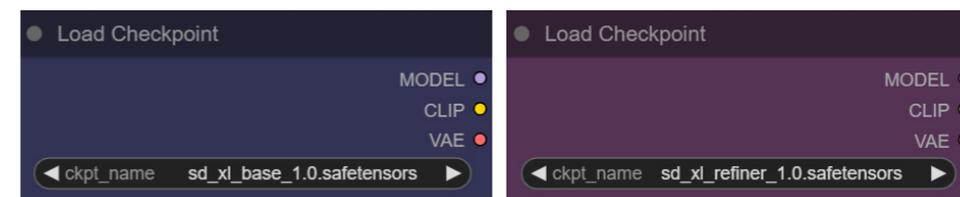


Fig 49. Funcionamiento Stable Diffusion 1

En segundo lugar, es necesario introducir un texto que servirá como prompt, para ello se utiliza un nodo de introducción de texto. Stable Diffusion tiene permite la introducción de un prompt positivo y uno negativo. En el nodo que servirá como prompt positivo, se ha de escribir aquello que se busca generar, mientras que en el negativo, es posible introducir aquellas cualidades que no se buscan en la imagen generada.



Fig 48. Funcionamiento Stable Diffusion 2

Este texto se procesa mediante un programa de procesamiento del lenguaje natural (PLN) para que sea procesable por los modelos de generación. Este se denomina “encoder” ya que codifica el prompt introducido para poder compararlo con la base de datos del entrenamiento.

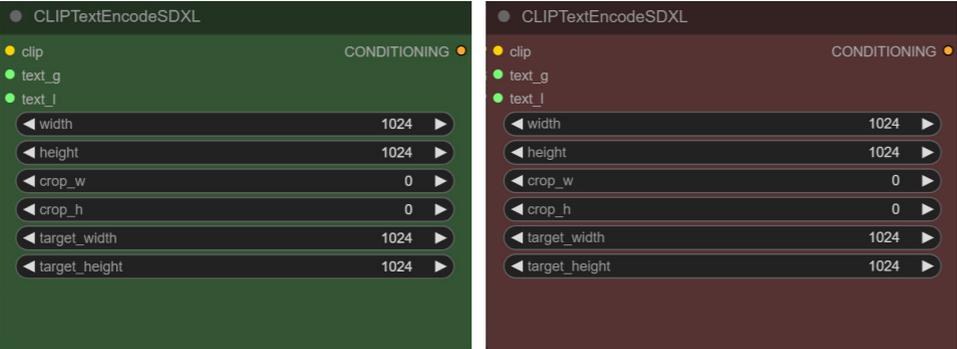


Fig 49. Funcionamiento Stable Diffusion 3

Por otro lado, es necesario introducir como input una imagen vacía de la resolución deseada, seleccionando de la misma manera, el número de divisiones dentro de la misma que definirán el marco de las imágenes generadas posteriormente.

En lugar de una imagen vacía, el programa permite de la misma manera introducir una imagen de referencia la cual tomará de referencia a la hora de generar los resultados.

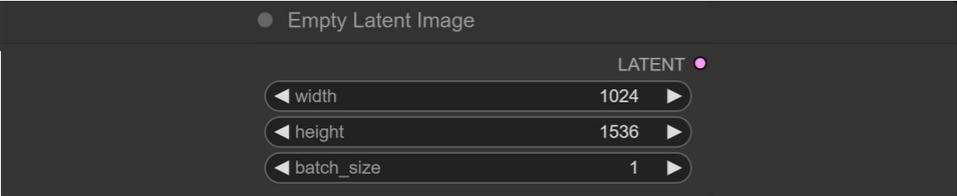


Fig 50. Funcionamiento Stable Diffusion 4

Todos estos datos junto con los parámetros pertinentes como la semilla de ruido aleatorio, el programador o los pasos de reducción de ruido a realizar, se han de introducir en un nodo denominado “KSampler”, en este punto se produce la difusión dando lugar a una serie de datos comprimidos que contienen las imágenes generadas.

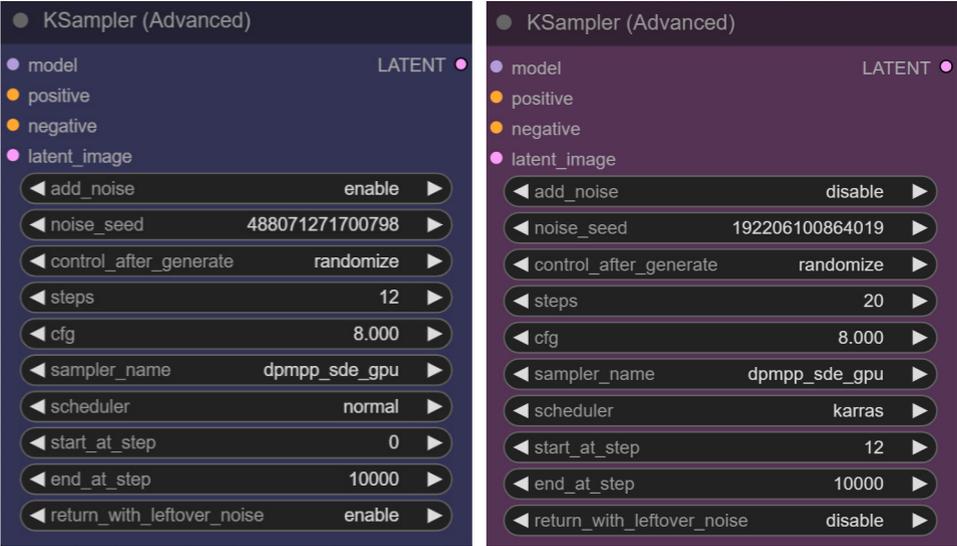


Fig 51. Funcionamiento Stable Diffusion 5

Para descomprimir dichos datos y transformarlos a un espacio en píxeles de la resolución deseada, es necesario un decodificador (decoder).

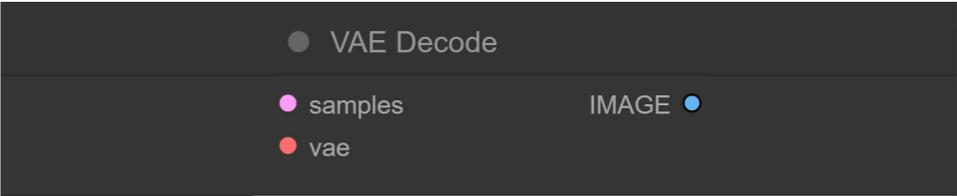


Fig 52. Funcionamiento Stable Diffusion 6

El proceso de generación comienza con el procesamiento de los datos introducidos para el modelo base, al finalizar este primer tramo, el programa permite obtener una previsualización de la imagen de baja resolución.

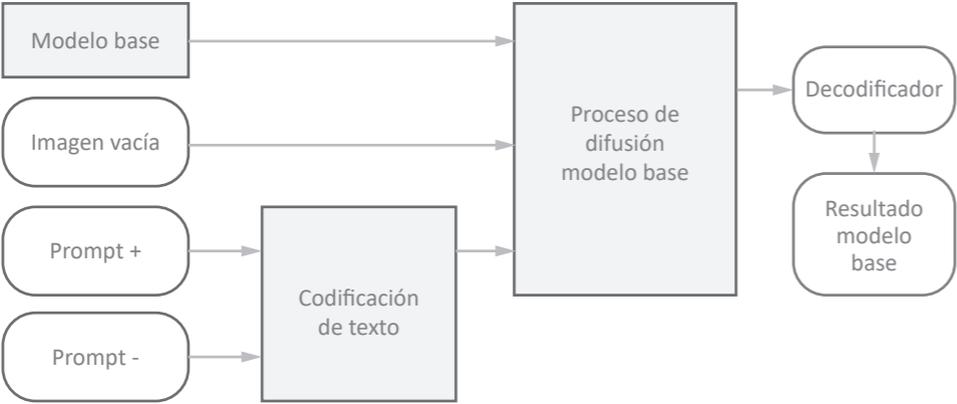


Fig 53. Esquema de funcionamiento

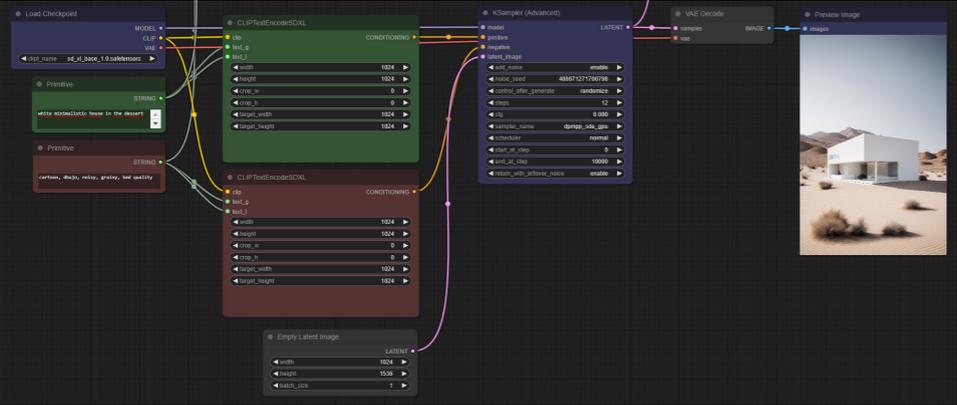


Fig 54. Funcionamiento Stable Diffusion 7



Fig 55. Resultado Stable Diffusion modelo base

Al finalizar la generación del modelo base, comienza la generación más detallada del modelo refinador. Este permite una previsualización de la imagen final. De no estar conforme con el resultado, es posible variar los resultados de ambos procesos o únicamente del refinador modificando el prompt y volviendo a ejecutar el programa.

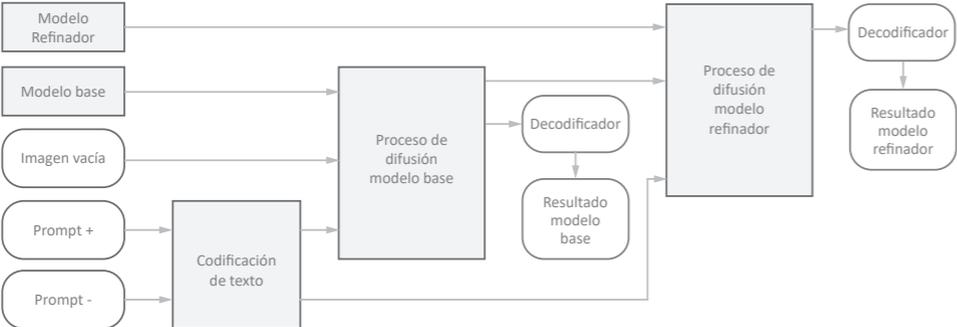


Fig 56. Esquema de funcionamiento 2

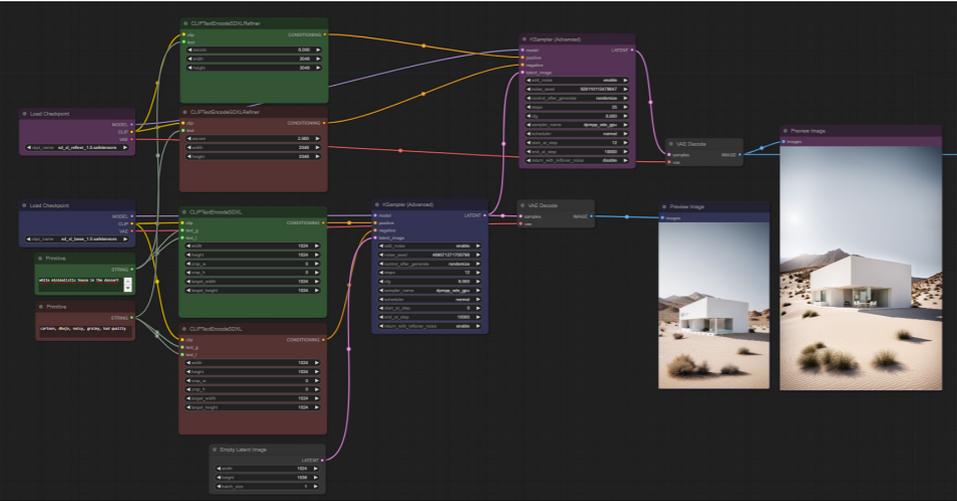


Fig 57. Funcionamiento Stable Diffusion 8



Fig 58. Resultado Stable Diffusion modelo refinador

Por último, gracias a esta interfaz, es posible añadir un modelo de escalado para obtener la imagen generada con una mayor resolución multiplicando esta por cuatro (4096x6194), obteniendo así una imagen de alta calidad que, siendo de mayor resolución, podría ser ajustada mediante programas de edición de imagen, como Adobe Photoshop.

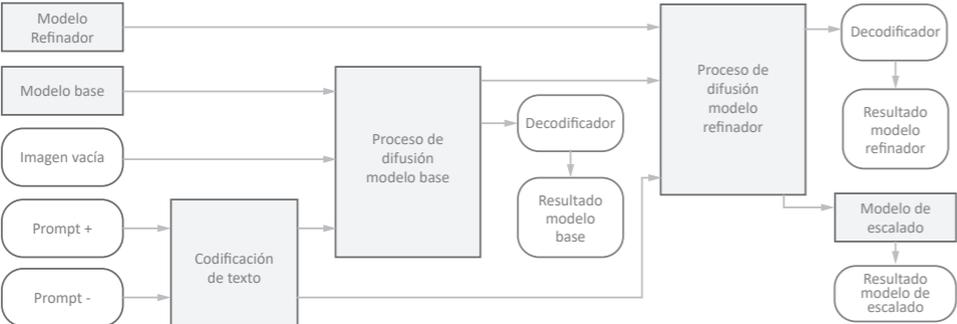


Fig 59. Esquema de funcionamiento 3

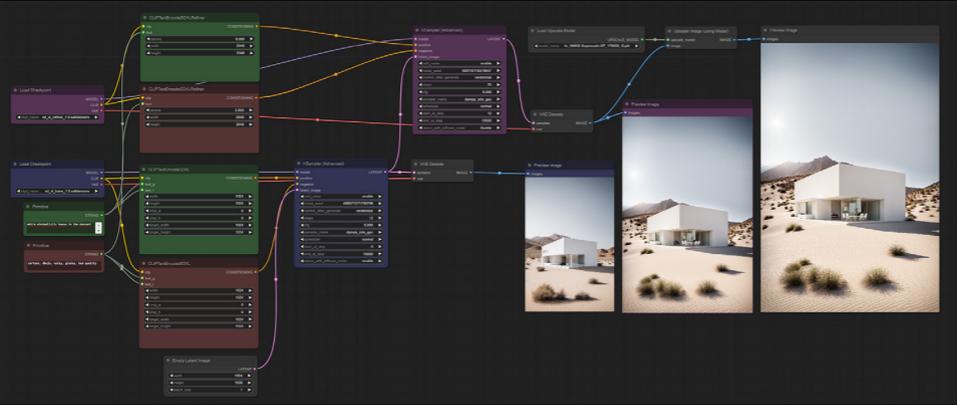


Fig 60. Funcionamiento Stable Diffusion 9



Fig 61. Resultado Stable Diffusion modelo de escalado

4. Comparativa de resultados

Para realizar la elección del programa que servirá como herramienta para la generación de la referencia para el proyecto, se realizarán una serie de pruebas a modo de comparación, que servirán para evaluar la capacidad de los distintos modelos de generación de imágenes seleccionados a la hora de generar elementos relacionados con la arquitectura. Como conjuntos que se consideran de interés para la generación de infografías arquitectónicas, se evaluará la capacidad de generación de los mismos en materia de anatomía, diseño de producto y arquitectura, en su expresión formal exterior e interior.

4.1 Análisis comparativo

Anatomía

Dall-e

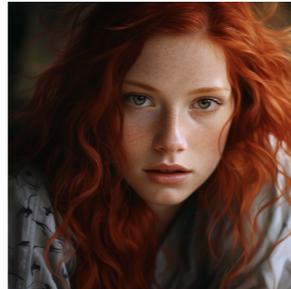
Midjourney

Stable Diffusion

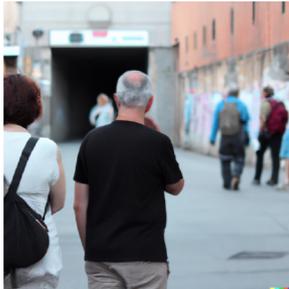
Photography of a human hand.



Realistic closeup portrait of a girl with red hair.



Realistic photography of people walking in the street.



Diseño

Dall-e

Midjourney

Stable Diffusion

Closeup high quality photography of a minimalistic concrete vase.



Gray and white minimalistic high-end chair, professional product design photography.



High quality photography of a minimalistic lamp.



Arquitectura exterior

Dall-e

Midjourney

Stable Diffusion

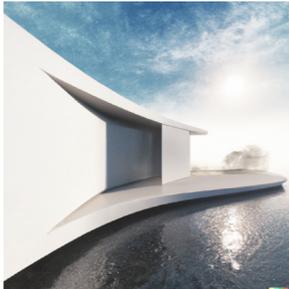
Minimal concrete house



Modern concrete house on a cliff in the mountains.



Curve minimalistic white house over a lake.



Arquitectura interior

Dall-e

Midjourney

Stable Diffusion

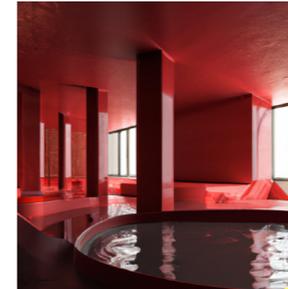
Timber lattice interior design space filtered light.



Brick structure pavilion inside photography.



Red concrete thermal spa interior design water space dim light.



Dall-e

A pesar de ser una de los programas de generación de imágenes con inteligencia artificial pioneros en el mercado, dall-e no ha continuado desarrollándose. Desde el lanzamiento de su segunda versión en septiembre de 2022, esta inteligencia artificial no ha sido reentrenada.

Este programa es capaz de comprender correctamente los textos introducidos. Aun así, en diversas ocasiones, ha sido necesario iterar los resultados más de una vez para conseguir un resultado aceptable. Como conclusión, este programa ofrece resultados aceptables, aunque presenta carencias graves en aspectos como la composición, el encuadre y la iluminación.

Midjourney

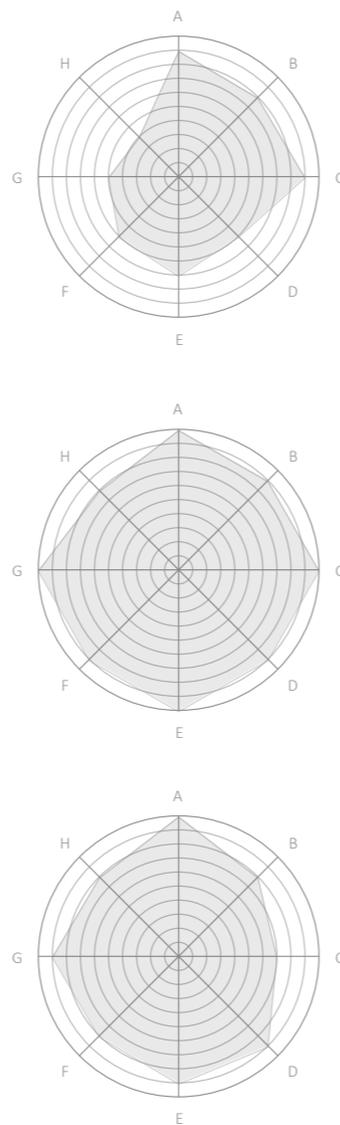
Las muestras obtenidas por la versión 5.2 de midjourney, en su mayoría destacan por su calidad. Esta inteligencia artificial se posiciona en la actualidad como una de las mejores alternativas a la hora de generar imágenes de diversas índoles, teniendo como punto negativo su elevado precio.

Teniendo en cuenta los resultados obtenidos, este programa destaca en aspectos como: comprensión, iluminación y encuadre ofreciendo imágenes de una gran calidad y creatividad, siendo esta una alternativa óptima para la generación de referencias en los campos del diseño y la arquitectura.

Stable Diffusion

Stable diffusion, en su última actualización, ha conseguido mejorar en gran medida la calidad de los resultados que ofrece. Estos resultados, junto al hecho de que sea un programa gratuito, lo convierten en una alternativa de alto rendimiento de los principales programas de pago del mercado.

En referencia a las muestras extraídas de este programa, se observa que estas presentan una gran precisión en aspectos como la creatividad, la composición, comprensión de los textos e iluminación de las imágenes generadas.



4.2 Conclusiones del análisis

Todos los programas de generación de imágenes con inteligencia artificial empleados para este análisis comparativo han generado resultados muy positivos poniendo en evidencia las altas capacidades de estas nuevas tecnologías. Este tipo de programas se encuentran en una fase muy primitiva en la actualidad, aun así, estos ofrecen resultados de gran calidad y han mejorado mucho en un corto periodo de tiempo.

Entre los tres modelos comparados, dall-e, por su falta de desarrollo, se encuentra en un nivel inferior a sus principales competidores, quedando descartada como el modelo a emplear para el desarrollo de este proyecto. Por otro lado, tanto midjourney como stable diffusion han generado resultados de alta calidad y precisión posicionándose como dos de las mejores alternativas actuales para la generación de imágenes con inteligencia artificial. Entre ellas, se pueden apreciar ligeras diferencias en cuanto a la composición e iluminación de las muestras que producen, siendo midjourney la que más destaca en estos aspectos.

Como conclusión, conviene subrayar que, tanto midjourney como stable diffusion serían buenas alternativas para la realización de los próximos apartados de este trabajo, no obstante, la capacidad de iteración y optimizado de los resultados de midjourney junto a la leve superioridad en cuanto a la calidad de las muestras que genera, hacen que esta inteligencia artificial sea la más adecuada para el desarrollo de este trabajo .

Fig 62-64. Gráficas comparativas, Pág. 72

- A Comprensión
- B Realismo
- C Lógica
- D Creatividad
- E Calidad de imagen
- F Composición
- G Iluminación
- H Encuadre

5. Desarrollo del proyecto

Como resultado del análisis comparativo, se concluye en la utilización de midjourney como programa de generación de imágenes de referencia. Para comenzar, el objetivo a seguir será el de extraer un pabellón efímero de madera con una complejidad formal que lo dote de interés a la hora de realizar el modelado. Tras realizar diversas pruebas, el resultado más aproximado al objetivo se obtiene mediante la introducción del prompt: "timber minimalistic parametric pavilion". Durante el proceso de obtención de la muestra, se realizan diversas modificaciones en el mismo, como añadir descripciones formales en referencia al entorno, el tipo de luz y la calidad de la imagen resultando en el prompt final: "high quality photography of a timber minimalistic parametric pavilion, water environment, blue hour light".

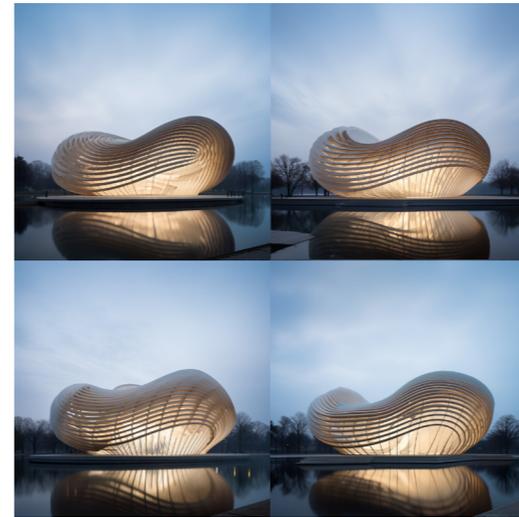
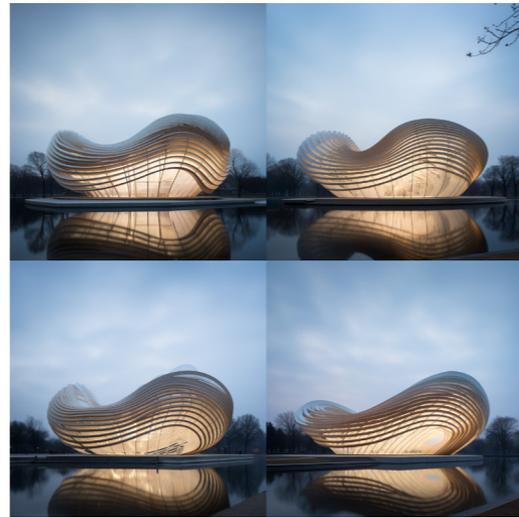
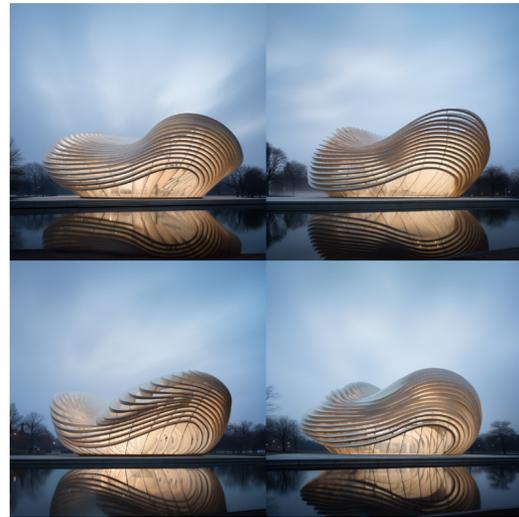
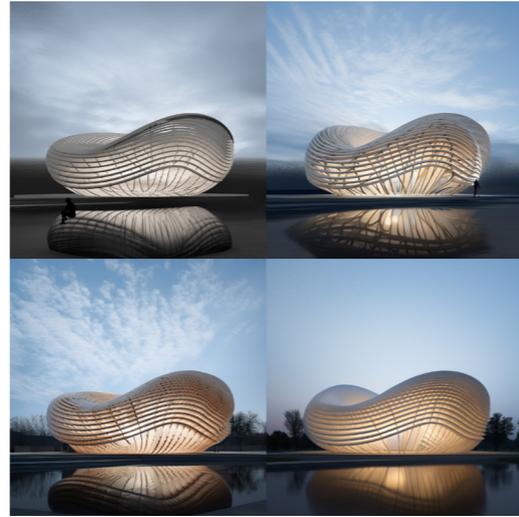
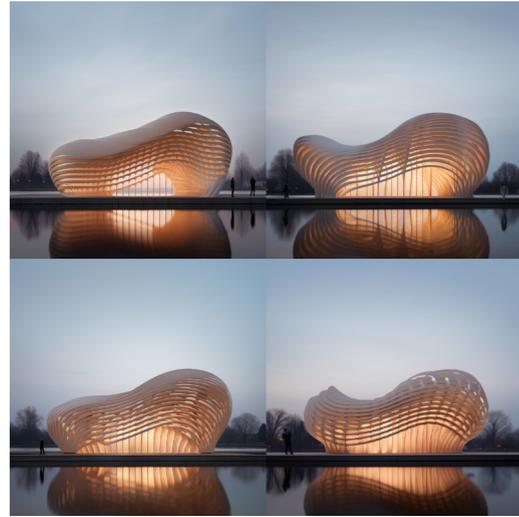


Fig 65-80. Obtención de la muestra, Págs. 78-81

5.2 Racionalización del objeto arquitectónico

Como resultado del proceso de iteración, se obtiene una referencia conceptual de interés. La imagen obtenida, muestra una construcción de madera formada por una serie de costillas verticales arriostradas por lamas horizontales curvas que abrazan el pabellón conformando su volumetría en toda su superficie. Dichas lamas, presentan una separación entre sí que confieren al conjunto de permeabilidad a la luz y al aire. Este elemento arquitectónico se encuentra presentado sobre un podio que irrumpe en la superficie de una lámina de agua, encontrándose situado en una zona aparentemente ajardinada.

Este pabellón presenta incongruencias formales que alejan al mismo de la realidad y dificultan su concepción como un objeto construido. Para este caso de estudio, se emplearán herramientas de modelado paramétrico para tratar de conferir a este objeto de una lógica constructiva y formal que permita su desarrollo posterior. Como resultado final se obtendrá una infografía 3D mediante el uso de programas específicos, como analogía del objeto construido, tratando de pasar, en la medida de lo posible, por las distintas fases de un proyecto arquitectónico.

Fig 81. Resultado del proceso de iteración, Pág 83



5.2.1 Proceso de modelado

El proceso de modelado del pabellón comienza con la introducción en Rhinoceros 3D de una curva base dibujada a mano que representaría la planta del pabellón. Esta se introduce en Grasshopper 3D partiendo de la misma para la realización de una serie de transformaciones asistidas por algoritmos.

En primer lugar, se aplica un desplazamiento vertical a la curva de base haciendo uso de una serie. Esta curva se copiará verticalmente tantas veces como iteraciones se definan para dicha serie y a una distancia n entre las mismas. Como resultado obtendremos una serie de curvas equidistantes.

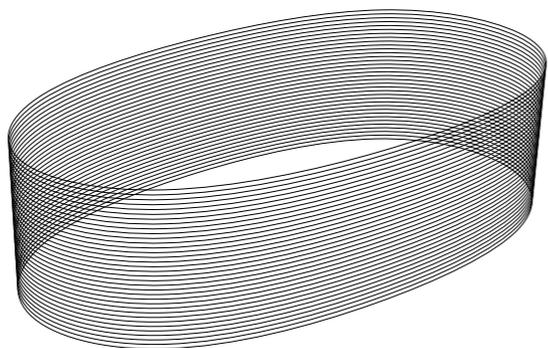
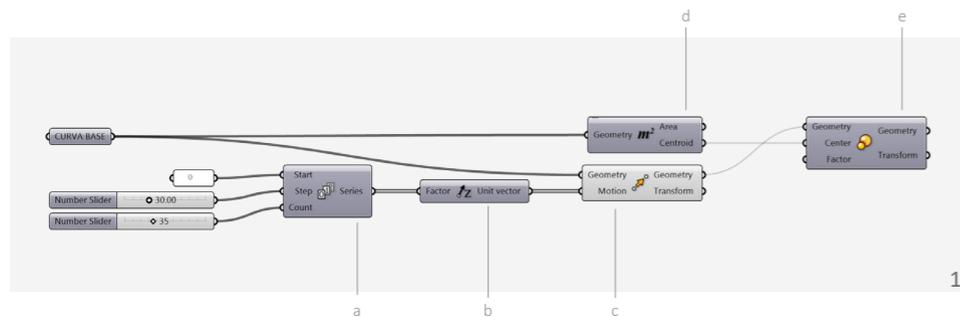


Fig 82. Proceso de modelado 1



A continuación, se aplica una transformación de escalado de las curvas provenientes del proceso anterior. El factor de escalado utilizado para la transformación de las mismas proviene de una distribución no lineal de los valores asociados a cada curva mediante una curva de Bézier. De esta manera, se genera una geometría aproximadamente toroidal que definirá el contorno del resultado final.

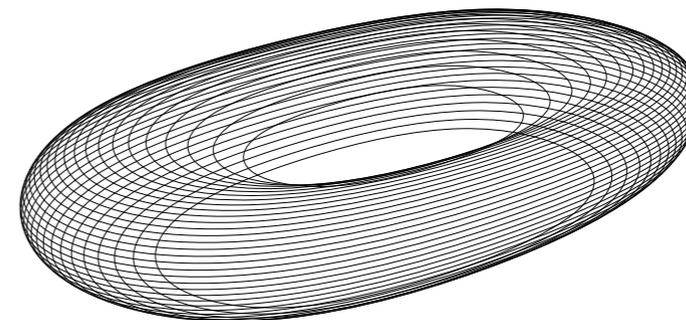
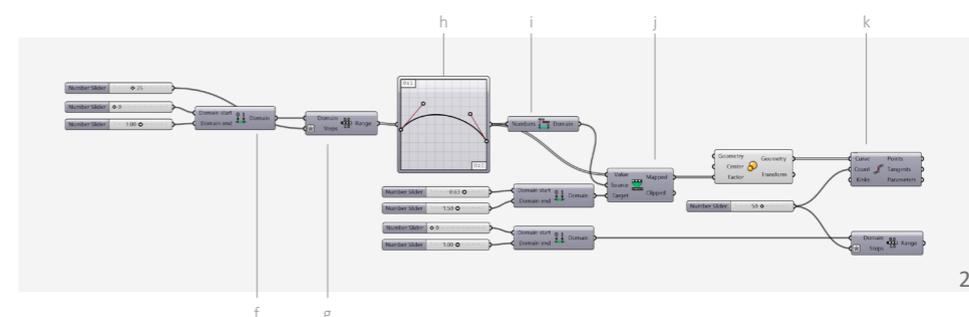


Fig 83. Proceso de modelado 2



Cada una de las curvas que forman esta geometría se divide en una serie de puntos que posteriormente se unirán consecutivamente para formar una primera aproximación de las costillas verticales que sustentarán el pabellón. En este momento todas las costillas tienen la misma altura, lo cual será modificado mediante la aplicación de dos transformaciones de escalado.

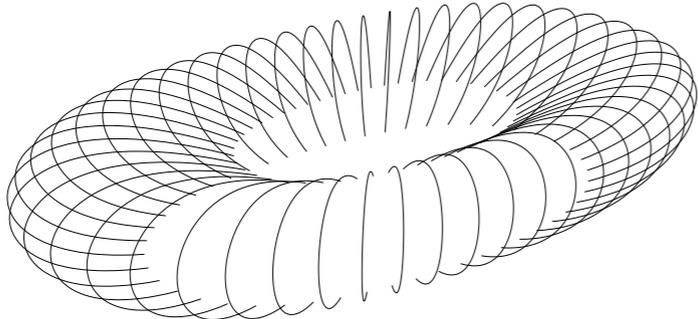
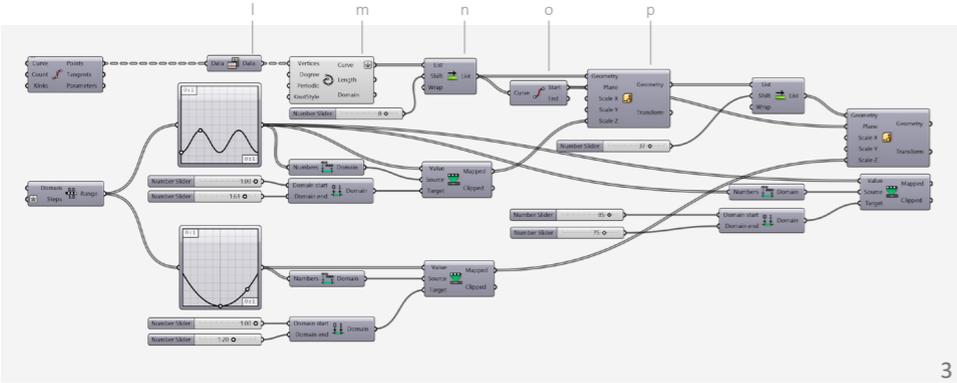


Fig 84. Proceso de modelado 3



3

Para la aplicación de un factor de escalado a cada una de las curvas se emplea, en primer lugar, una distribución sinusoidal de los valores produciendo mayores resultados para los extremos y menores para las partes centrales. En segundo lugar, se aplica un escalado con una distribución parabólica de los factores asociados a cada curva consiguiendo, de esta manera, la elevación de uno de los dos extremos de la geometría.

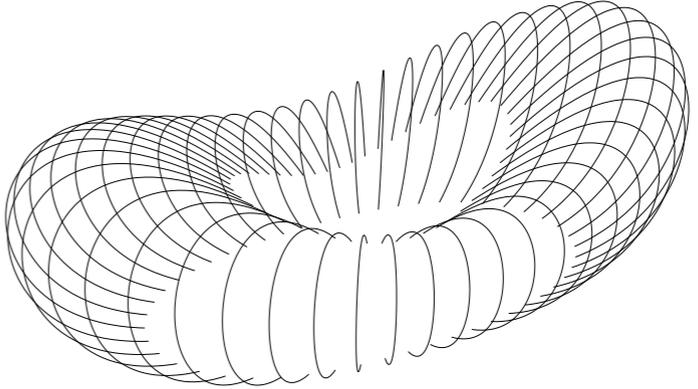
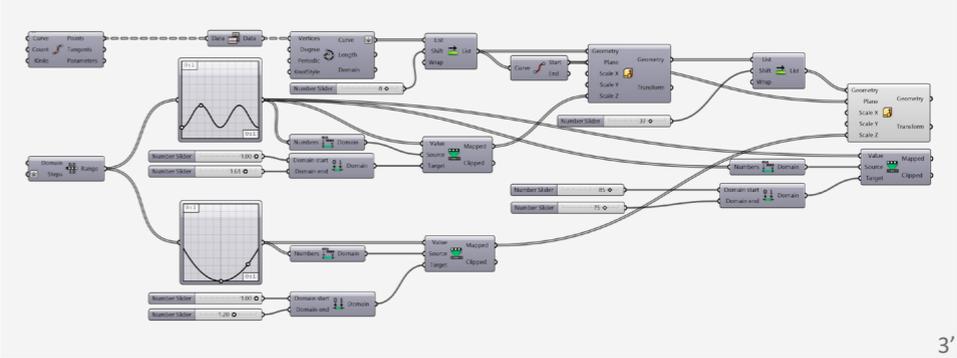


Fig 85. Proceso de modelado 4



3'

Al realizar la unión de dichos puntos horizontalmente, se consigue la curvatura horizontal de las lamas que recubrirán el pabellón, ya que dicha unión se realiza un número de veces no mayor a la cantidad de divisiones de las curvas de menor longitud.

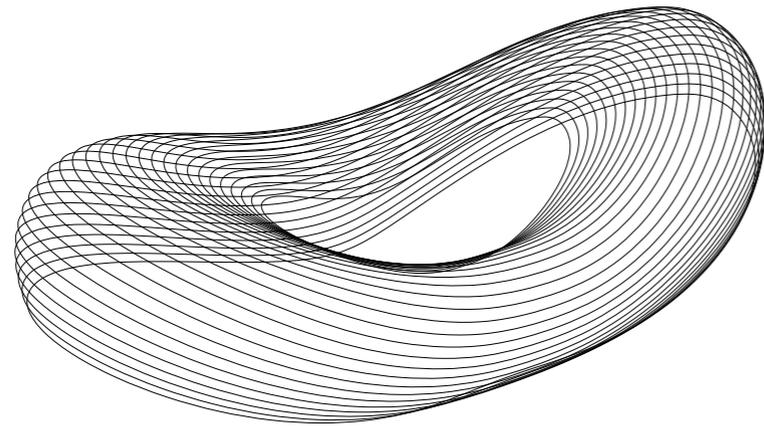
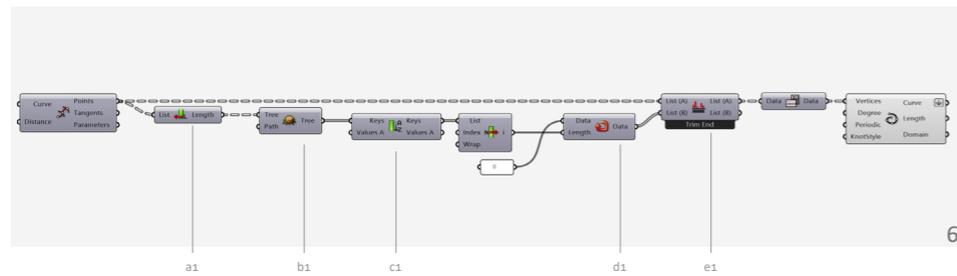


Fig 88. Proceso de modelado 7



Posteriormente, se realiza un desfase de las curvas hacia el exterior y una extrusión de la superficie generada por la combinación de ambas, consiguiendo el resultado final de los elementos que arriostarán la estructura horizontalmente.

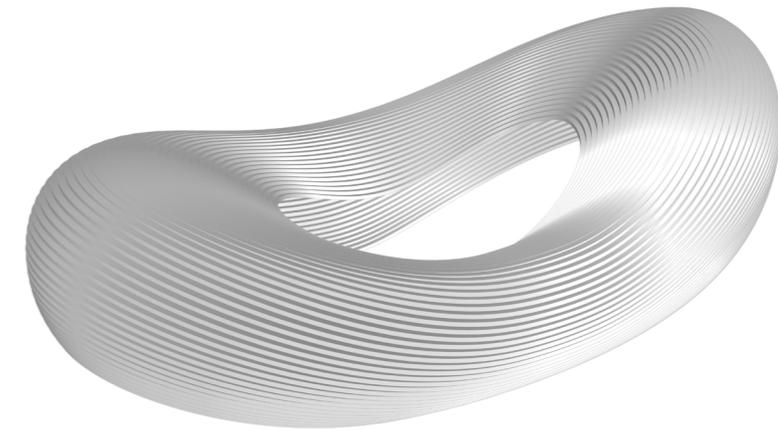
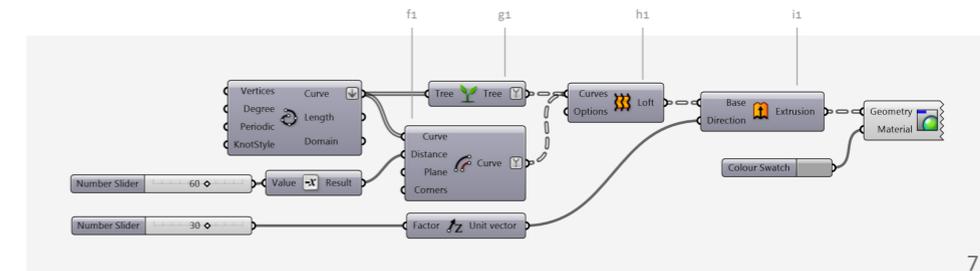


Fig 89. Proceso de modelado 8



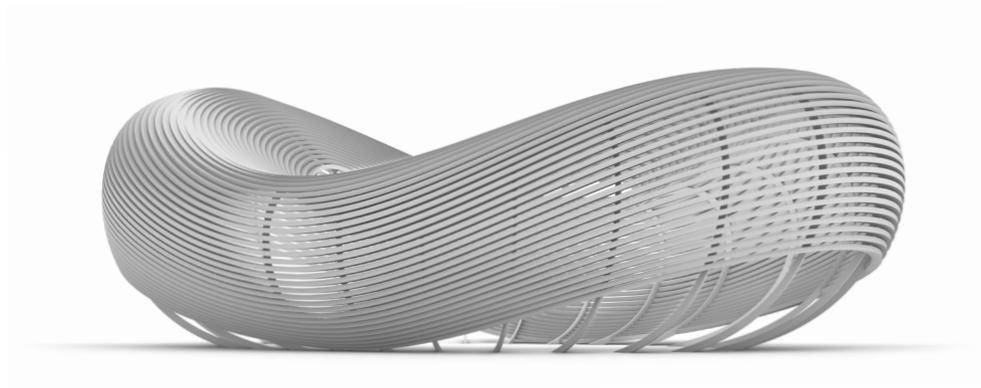


Fig 90. Resultado del modelado

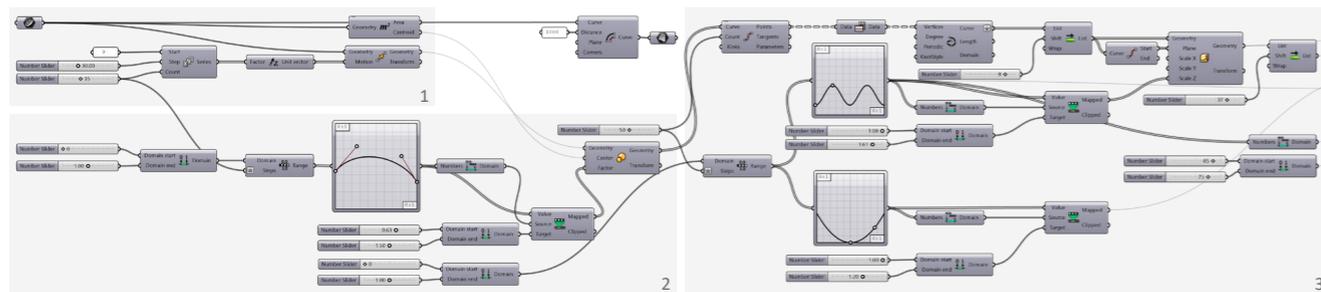
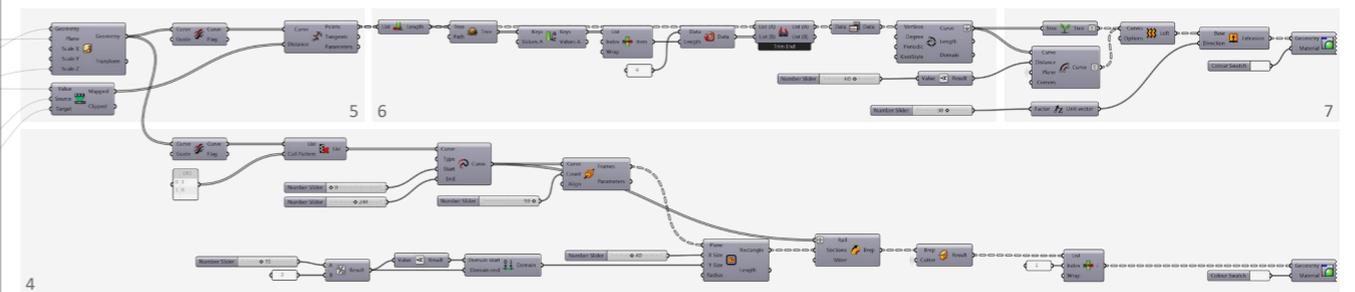


Fig 91. Código de modelado

- | | | | | |
|----------|-----------------|----------------|-------------------|------------------|
| a serie | h graph mapper | o end points | v sweep 1 | c1 sort keys |
| b unit z | i bounds | p scale NU | w split brep | d1 repeat data |
| c move | j remap numbers | q flip curve | x list item | e1 shortest list |
| d area | k divide curve | r cull pattern | y custom preview | f1 offset curve |
| e scale | l flip matrix | s extend curve | z divide distance | g1 graft tree |
| f domain | m interpolate | t perp frames | a1 list length | h1 loft |
| g range | n shift list | u rectangle | b1 flatten tree | i1 extrude |



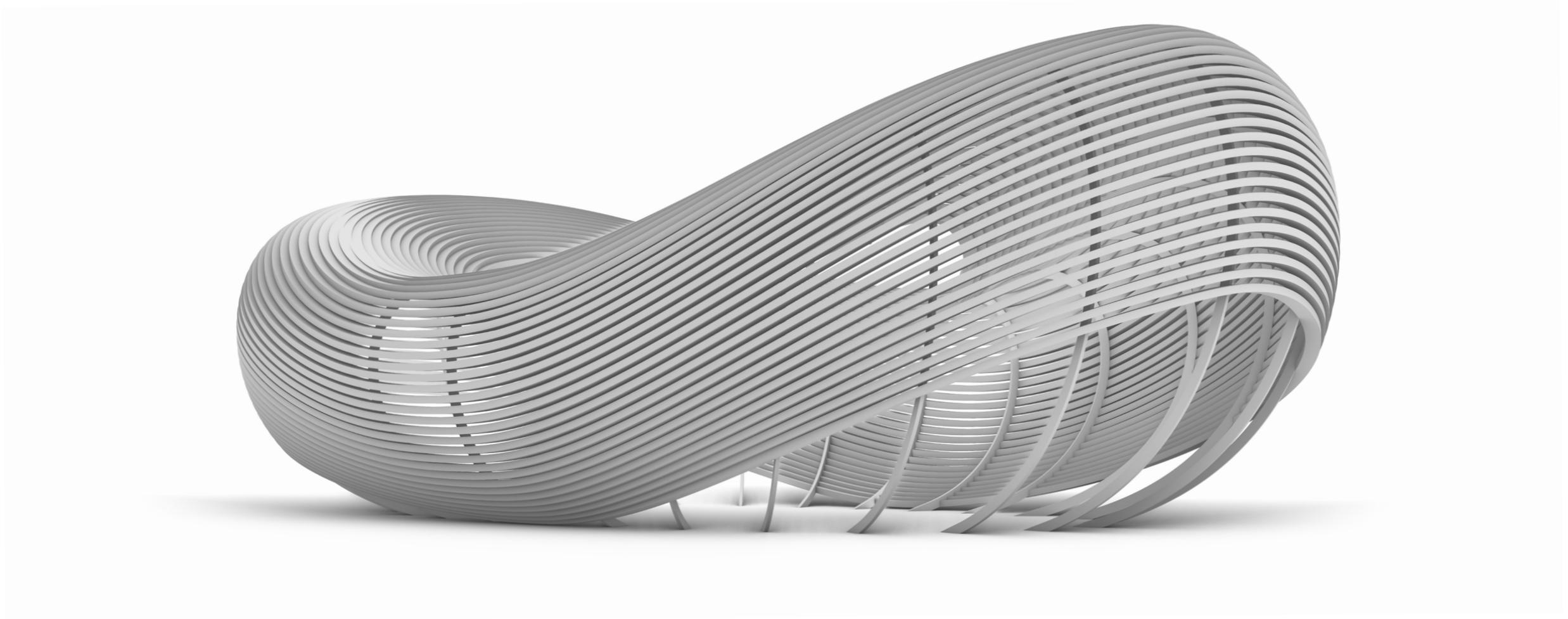


Fig 92. Resultado del modelado 2

5.3 Resultado final

Una vez culminado el proceso de modelado del elemento arquitectónico, se exporta el modelo a una plataforma de visualización especializada, en este caso, Autodesk 3Ds Max. En esta etapa, se concentra la atención en tres aspectos clave: materialidad, entorno e iluminación. La materialidad implica la aplicación de texturas y materiales realistas a las superficies del modelo. Luego, se crea el entorno circundante, incorporando elementos como árboles, para contextualizar el diseño. Por último, la iluminación desempeña un papel fundamental en la creación de una infografía convincente, configurándose las fuentes de iluminación natural y artificial tratando de lograr una atmósfera similar a la de la imagen de referencia.

Una vez que estos tres aspectos se han realizado en el programa, se procede a la etapa de renderización, que consiste en generar una imagen final de alta resolución a partir del modelo 3D y las configuraciones previamente establecidas, dejando que se realicen los procesos necesarios el tiempo suficiente para obtener una infografía de alta calidad.

Fig 93. Resultado final



6. Conclusiones

Es un hecho que las nuevas tecnologías resultan disruptivas y muchas veces pueden llegar a ser percibidas como una amenaza, el periodo que estamos pudiendo presenciar, no es distinto a hechos históricos como la llegada de la máquina a las fábricas y, más recientemente, los ordenadores a la mayoría de los puestos de trabajo de profesiones como la que nos compete. Si bien es cierto que los avances tecnológicos pueden llegar a restar valor a ciertas tareas realizadas por el ser humano, estos pueden llevar a un aumento muy elevado de la productividad consiguiendo resultados de una alta precisión a una velocidad inconcebible hasta el momento. Dichos avances promueven un modelo competitivo y de constante cambio en busca de un continuo avance, mejora y optimización de los recursos. Este proceso, se encuentra en concordancia con las teorías de la evolución presentes en la naturaleza y resulta imprescindible para el desarrollo natural del ser humano.

Por otra parte, cabe destacar que este trabajo no pretende suscitar a las nuevas generaciones de profesionales del sector creativo a la utilización indiscriminada e inmoral de estas tecnologías, sino que trata de ser una exploración formal de una de las herramientas que, a día de hoy, los arquitectos y diseñadores tenemos a nuestra disposición. Estas herramientas, al igual que la mayoría de elementos de nuestro entorno, pueden ser empleadas para bien y para mal, quedando a juicio y conciencia del individuo el uso que finalmente se les dé. Así pues, resulta de gran importancia realizar un llamamiento a la protección de la autoría y a una concienciación de la población en términos de conseguir comportamientos responsables para que, en el momento en que el uso de este tipo de tecnologías sea generalizado, la mayoría de las personas tengan un criterio suficiente para decidir como las emplean de una manera consciente.

“Si uno no cambia, no evoluciona y termina por dejar de pensar.” (Rem Koolhaas, 1944-)

“Cada nueva situación requiere una nueva arquitectura.” (Jean Nouvel, 1945-)

“Las buenas ideas vienen de todas partes. Es más importante reconocer una buena idea que crearla.” (Jeanne Gang, 1964-)

Referencias bibliográficas

1. Castro Pena, M. L., et al, 2021. *Artificial intelligence applied to conceptual design. A review of its use in architecture*, Automation in Construction 124 103550.
2. Russell, S., Norvig, P., 1995. *Artificial intelligence: A modern approach*. Upper Saddle River, New Jersey: Prentice Hall.
3. López Takeyas, B., 2007. *Introducción a la inteligencia artificial*, Instituto Tecnológico de Nuevo Laredo Nuevo Laredo, Tamps. México. Reforma Sur.
4. <https://vfarquitecto.com/buildtech-arquitectura-impulsada-por-inteligencia-artificial-y-aprendizaje-automatico/>
5. Senovilla tejedor, A., Álvarez alonso, M., 2020. *Inteligencia artificial y aprendizaje automático en la arquitectura*, Universidad politécnica de madrid. Madrid.
6. Nishanta, R., Kennedy, M. , Corbett, J., 2020. "Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda." *International Journal of Information Management* 53.
7. Wang, Z., Srinivasan, R. S., 2017. "A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models". *Renewable and Sustainable Energy Reviews*, Volumen 75.
8. Zhou, L., et al., 2018. Emergency decision making for natural disasters: An overview. *International Journal of Disaster Risk Reduction*. Volumen 27.
9. Demirci, M., Unes, F. y Korlu, S., 2019. Modeling of groundwater level using artificial intelligence techniques: A case study of Reyhanli region in Turkey. *Applied Ecology and Environmental Research*, 17(2), 2651-2663.
10. <https://datascientest.com/es/deep-learning-definicion>
11. <https://www.iartificial.net/redes-neuronales-generativas-adversarias-gans/>
12. <https://hashdork.com/es/diffusion-model-vs-gan/>
13. <https://theaisummer.com/diffusion-models/>
14. <https://impulsatek.com/los-modelos-de-difusion-claramente-explicados/>
15. <https://lilianweng.github.io/posts/2021-07-11-diffusion-models/>
16. Yu, J., et al., 2023. *CgT-GAN: CLIP-guided Text GAN for Image Captioning*. *arXiv:2308.12045v1*, 23.
17. <https://openai.com/research/clip>
18. Witteveen, S., Andrews, M., 2022. *Investigating Prompt Engineering in Diffusion Models*. *arXiv:2211.15462v1* [cs.CV].
19. Tedeschi, A., 2014. *AAD_Algorithms aided design*. Le Penseur Publisher.
20. García Alvarado, R., Lyon Gottlieb, A., 2013. "Diseño paramétrico en Arquitectura: método, técnicas y aplicaciones", *Arquisur Revista*, Año 3, Vol 3.
21. https://es.wikipedia.org/wiki/Visualización_de_Arquitectura
22. <https://es.cointelegraph.com/news/history-and-mission-of-openai-the-company-behind-chatgpt>
23. <https://openai.com/blog/planning-for-agi-and-beyond>
24. <https://noticias.ai/midjourney/>
25. <https://waxy.org/2022/08/exploring-12-million-of-the-images-used-to-train-stable-diffusions-image-generator/>
26. https://es.wikipedia.org/wiki/Stable_Diffusion
27. Schuhmann, C., et al., 2022. LAION-5B: An open large-scale dataset for training next generation image-text models.
28. <https://stablediffusionxl.com/>

Bibliografía

1. Castro Pena, M. L., et al, 2021. *Artificial intelligence applied to conceptual design. A review of its use in architecture*, Automation in Construction 124 103550.
2. Demirci, M., Unes, F. y Korlu, S., 2019. Modeling of groundwater level using artificial intelligence techniques: A case study of Reyhanli region in Turkey. *Applied Ecology and Environmental Research*, 17(2), 2651-2663.
3. Fu, T., 2023. MJ Studio 3. Masterclass.
4. García Alvarado, R., Lyon Gottlieb, A., 2013. "Diseño paramétrico en Arquitectura: método, técnicas y aplicaciones", *Arquisur Revista*, Año 3, Vol 3.
5. López Takeyas, B., 2007. *Introducción a la inteligencia artificial*, Instituto Tecnológico de Nuevo Laredo Nuevo Laredo, Tamps. México. Reforma Sur.
6. Nishanta, R., Kennedy, M. , Corbett, J., 2020. "Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda." *International Journal of Information Management* 53.
7. Russell, S., Norvig, P., 1995. *Artificial intelligence: A modern approach*. Upper Saddle River, New Jersey: Prentice Hall.
8. Schuhmann, C., et al., 2022. LAION-5B: An open large-scale dataset for training next generation image-text models.
9. Senovilla tejedor, A., Álvarez alonso, M., 2020. *Inteligencia artificial y aprendizaje automático en la arquitectura*, Universidad politécnica de madrid. Madrid.
10. Wang, Z., Srinivasan, R. S., 2017. "A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models". *Renewable and Sustainable Energy Reviews*, Volumen 75.
11. Witteveen, S., Andrews, M., 2022. *Investigating Prompt Engineering in Diffusion Models*. arXiv:2211.15462v1 [cs.CV].
12. Yu, J., et al., 2023. *CgT-GAN: CLIP-guided Text GAN for Image Captioning*. arXiv:2308.12045v1, 23.
13. Zhou, L., et al., 2018. Emergency decision making for natural disasters: An overview. *International Journal of Disaster Risk Reduction*. Volumen 27.

Websites

1. <https://vfarquitecto.com/buildtech-arquitectura-impulsada-por-inteligencia-artificial-y-aprendizaje-automatico/>
2. <https://datascientest.com/es/deep-learning-definicion>
3. <https://www.iartificial.net/redes-neuronales-generativas-adversarias-gans/>
4. <https://hashdork.com/es/diffusion-model-vs-gan/>
5. <https://theaisummer.com/diffusion-models/>
6. <https://impulsatek.com/los-modelos-de-difusion-claramente-explicados/>
7. <https://lilianweng.github.io/posts/2021-07-11-diffusion-models/>
8. https://es.wikipedia.org/wiki/Visualización_de_Arquitectura
9. <https://es.cointelegraph.com/news/history-and-mission-of-openai-the-company-behind-chatgpt>
10. <https://openai.com/blog/planning-for-agi-and-beyond>
11. <https://noticias.ai/midjourney/>
12. <https://waxy.org/2022/08/exploring-12-million-of-the-images-used-to-train-stable-diffusions-image-generator/>
13. https://es.wikipedia.org/wiki/Stable_Diffusion
14. <https://stablediffusionxl.com/>

Tabla de figuras

Figura 01 Redes neuronales. Elaboración propia.
Figura 02 Redes generativas adversarias. Elaboración propia.
Figura 03 Modelos de difusión. <https://saurabhraj5162.medium.com/diving-into-diffusion-models-e452a299f343>
Figura 04 OpenAi Clip. <https://openai.com/>
Figura 05 Ingeniería de prompts. MJ Studio 3, Tim Fu.
Figura 06 Centro Heydar aliyev, Zaha Hadid. <https://www.archdaily.cl/cl/02-203430/feliz-cumpleanos-zaha-hadid>
Figura 07 Serpentine Gallery, BIG. <https://www.archdaily.cl/cl/tag/serpentine-gallery>
Figura 08 V & A Dundee, Kengo Kuma. <https://www.floornature.es/kengo-kuma-v-dundee-escocia-14056/>
Figura 09 Algoritmo Grasshopper. Elaboración propia.
Figura 10 Diseño paramétrico. Elaboración propia.
Figura 11 Programas de modelado paramétrico. <https://pngio.com/images/png-x1712293.html>
Figura 12 Programas de visualización 3D. https://favpng.com/png_view/3ds-max-v-ray-rendering-node-autodesk-maya-rhinoceros-3d-png/1UviAtvh
Figura 13 Logo Dall-e. <https://freelogopng.com/dall-e-logo-png>
Figura 14 Resultado Dall-e. Elaboración propia.
Figuras 15-22 Funcionamiento Dall-e. Elaboración propia.
Figura 23 Programa de escalado. <https://replicate.com/nightmareai/real-esrgan/>
Figura 24 Resultado escalado. Elaboración propia.
Figura 25 Logo Midjourney. <https://forobeta.com/ia/midjourney.3/>
Figura 26 Resultado Midjourney. Elaboración propia.
Figuras 27-44 Funcionamiento Midjourney. Elaboración propia.
Figura 45 Escalado Midjourney. Elaboración propia.
Figura 46 Logo Stable Diffusion. <https://tngd.sergeswin.com/es/12724/>
Figura 47 Resultado Stable Diffusion. Elaboración propia.
Figuras 48-52 Funcionamiento Stable Diffusion. Elaboración propia.
Figura 53 Esquema de funcionamiento Stable Diffusion 0. Elaboración propia.

Figura 54 Funcionamiento Stable Diffusion. Elaboración propia.
Figura 55 Resultado modelo base Stable Diffusion. Elaboración propia.
Figura 56 Esquema de funcionamiento Stable Diffusion 1. Elaboración propia.
Figura 57 Funcionamiento Stable Diffusion. Elaboración propia.
Figura 58 Resultado modelo refinador Stable Diffusion. Elaboración propia.
Figura 59 Esquema de funcionamiento Stable Diffusion 2. Elaboración propia.
Figura 60 Funcionamiento Stable Diffusion. Elaboración propia.
Figura 61 Resultado modelo escalado Stable Diffusion. Elaboración propia.
Figuras 62-64 Gráficas comparativas. Elaboración propia.
Figuras 65-80 Obtención de la muestra. Elaboración propia.
Figura 81 Resultado del proceso de iteración. Elaboración propia.
Figuras 82-89 Proceso de modelado. Elaboración propia.
Figura 90 Resultado del modelado. Elaboración propia.
Figura 91 Código del modelado. Elaboración propia.
Figura 92 Resultado del modelado 2. Elaboración propia.
Figura 93 Resultado final. Elaboración propia.

Anexo

Catálogo de imágenes generadas con midjourney

Como evidencia de la capacidad de estas tecnologías, se presenta una recopilación de imágenes generadas a lo largo del desarrollo de este proyecto. Estas imágenes son el resultado tangible de la aplicación de la inteligencia artificial en el ámbito de la arquitectura, revelando diseños y conceptos de gran potencia.

