



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Politécnica Superior de Gandia

Modelos de Inteligencia Artificial en la generación de
imágenes: situación actual y análisis comparativo

Trabajo Fin de Grado

Grado en Comunicación Audiovisual

AUTOR/A: Miñana Sánchez, Eva

Tutor/a: Alonso López, Nàdia

CURSO ACADÉMICO: 2023/2024

RESUMEN

El presente Trabajo Final de Grado tiene como objetivo estudiar la situación actual de la Inteligencia Artificial Generativa en el ámbito social y audiovisual, y realizar una evaluación de los modelos de imagen estática creados por esta técnica de cuatro de las compañías más grandes del mercado en este sector actualmente, como son DALL·E, Midjourney, Leonardo AI y Stable Diffusion.

Para ello, se examinará la posición de esta tecnología en el panorama global que nos rodea, la evolución que sigue, su alcance e influencia y las repercusiones que tiene hoy en día en las áreas definidas. De este modo, se efectuará un análisis en detalle individual de cada uno de los modelos mencionados anteriormente, con sus respectivas funciones, dónde se presentarán sus beneficios y desventajas, así como sus similitudes y diferencias para entender cómo se han desarrollado y su forma de trabajo. Posteriormente, se expondrán en conjunto para entender la estrategia que emplean y sus correspondientes respuestas.

Cabe destacar que toda la información y datos recogidos en esta memoria alcanzan hasta junio de 2024. Dado que esta tecnología está en evolución constante, es probable que surjan nuevas funcionalidades más adelante que vayan más allá de lo expuesto en el presente Trabajo Final de Grado.

PALABRAS CLAVE

Inteligencia Artificial, DALL·E, Midjourney, Leonardo AI y Stable Diffusion.

ABSTRACT

The objective of this Final Degree Project is to study the current situation of Generative Artificial Intelligence in the social and audiovisual fields, and to evaluate the static image models created by this technique from four of the largest companies in the market in this sector today: DALL·E, Midjourney, Leonardo AI, and Stable Diffusion.

To achieve this, the position of this technology in the global landscape will be examined, including its evolution, scope, influence, and the repercussions it currently has in the defined areas. Thus, a detailed individual analysis of each of the aforementioned models will be conducted, including their respective functions, highlighting their benefits and disadvantages, as well as their similarities and differences to understand how they have developed and their working methods. Subsequently, they will be presented together to understand the strategies they employ and their corresponding responses.

It is important to note that all the information and data collected in this work is current up to June 2024. Since this technology is constantly evolving, it is likely that new functionalities will emerge beyond what is presented in this Final Degree Project.

KEY WORDS

Artificial Intelligence, DALL·E, Midjourney, Leonardo AI and Stable Diffusion.

ÍNDICE

1. INTRODUCCIÓN	6
1.1. Justificación del tema.....	6
1.2. Objetivos.....	7
1.3. Metodología.....	7
1.4. Etapas.....	9
2. LA IA GENERATIVA EN EL AUDIOVISUAL	9
2.1. Primeros años.....	9
2.2. Contexto histórico y social.....	11
2.3. Estructura y funcionamiento.....	13
2.4. La generación de imágenes.....	16
2.5. Modelos actuales.....	19
3. ANÁLISIS DE LOS MODELOS DE IA GENERATIVA	24
3.1. DALL·E.....	24
3.2. Midjourney.....	30
3.3. Leonardo AI.....	37
3.4. Stable Difussion.....	41
3.5. Comparativa.....	45
4. CONCLUSIONES	47
5. BIBLIOGRAFÍA	48
ANEXO: Relación del trabajo con los Objetivos de Desarrollo Sostenible de la Agenda 2030	

ÍNDICE DE FIGURAS

Figura 1. Estructura de una red neuronal profunda en <i>Deep Learning</i>	14
Figura 2. Arquitectura interna de una GAN.....	15
Figura 3. <i>Deepfake</i> del Papa Francisco con un abrigo de la firma Balenciaga	18
Figura 4. <i>Deepfake</i> de los expresidentes de los Estados Unidos, Barack Obama y Donald Trump, jugando un partido de baloncesto	19
Figura 5. Página web principal de ChatGPT donde se inicia sesión	25
Figura 6. Herramienta DALL·E preparada para su ejecución	25
Figura 7. Resultados de imagen estática generados por IA en DALL·E mediante un prompt de texto.....	26
Figura 8. Resultados de imagen estática generados por IA en DALL·E modificando el prompt original desde el cuadro de texto	27
Figura 9. Resultados de imagen estática generados por IA en DALL·E modificando el prompt desde la imagen original	27
Figura 10. Funciones predeterminadas para la modificación de imágenes en DALL·E	28
Figura 11. Resultados de imagen estática generados por IA en DALL·E modificando el prompt original ahora con un ratio 16:9 y un estilo abstracto	28
Figura 12. Funciones establecidas dentro de una imagen en DALL·E	29
Figura 13. Visualización del prompt completo en una imagen en DALL·E.....	29
Figura 14. Proceso para transformar una imagen en DALL·E con la herramienta seleccionar	29
Figura 15. Creación de un servidor en Discord	30
Figura 16. Menú de exploración de servidores de Discord	31
Figura 17. Proceso para implementar el bot Midjourney en un servidor personal de Discord (I).....	32
Figura 18. Proceso para implementar el bot Midjourney en un servidor personal de Discord (II).....	32
Figura 19. Resultados de imagen estática generados por IA en Midjourney mediante un <i>prompt</i> de texto.....	33
Figura 20. Resultados de imagen estática generados por IA en Midjourney modificando el <i>prompt</i> desde las funciones <i>Vary</i>	34
Figuras 21a y 21b. Resultados de imagen estática generados por IA en Midjourney desde las funciones <i>Vary suave</i> (izq.) y <i>Vary fuerte</i> (dcha.).....	35
Figura 22. Resultados de imagen estática generados por IA en Midjourney modificando el prompt original	35
Figura 23. Proceso para transformar una imagen en Midjourney con la herramienta <i>Vary Region</i>	36
Figura 24. Resultados de imagen estática generados por IA en Midjourney modificando el prompt original ahora con estilo basado en el s.XVII, pintura al óleo y con un ratio 3:2	36
Figura 25. Página web principal de Leonardo AI donde se inicia sesión.....	37
Figura 26. Menú principal de Leonardo AI con todas las herramientas disponibles....	37
Figura 27. Configuración preestablecida y resultados de imagen estática generados por IA en Leonardo AI mediante un <i>prompt</i> de texto con el motor <i>Leonardo Phoenix</i>	38
Figura 28. Visualización del <i>prompt</i> completo de una generación de imágenes en Leonardo AI	39

Figuras 29a y 29b. Resultados de imagen estática generados por IA en Leonardo AI mediante un <i>prompt</i> de texto con el motor <i>Lifelike Vision (izq.)</i> y <i>Stock Photography (dcha.)</i>	39
Figura 30. Resultados de imagen estática generados por IA en Leonardo AI modificando el <i>prompt</i> original con el motor <i>Stock Photography</i>	40
Figura 31. Resultados de imagen estática generados por IA en Leonardo AI modificando el <i>prompt</i> original con estilo basado en el s.XVII, pintura al óleo y con un ratio 16:9 (sin efecto en la respuesta)	40
Figura 32. Resultados de imagen estática generados por IA en <i>Leonardo AI</i> modificando el <i>prompt</i> original con el motor <i>Anime</i>	41
Figura 33. Página web principal de Dream Studio donde se inicia sesión	42
Figura 34. Interfaz de generación estática de Dream Studio. Extraído de Dream Studio	42
Figuras 35a y 35b. Resultados de imagen estática generados por IA en Dream Studio mediante un <i>prompt</i> de texto con los motores de Stable Difussion <i>Photographic (izq.)</i> y <i>Cinematic (dcha.)</i>	43
Figura 36. Resultados de imagen estática generados por IA en Dream Studio mediante un <i>prompt</i> de texto sin ningún motor de renderizado	43
Figura 37. Resultados de imagen estática generados por IA en Dream Studio modificando el <i>prompt</i> original sin ningún motor de renderizado	44
Figura 38. Proceso para transformar una imagen en Dream Studio con la herramienta seleccionar	44
Figura 39. Resultados de imagen estática generados por IA en Dream Studio modificando el <i>prompt</i> original ahora con estilo basado en el s.XVII, pintura al óleo y con un ratio 16:9	44

1. INTRODUCCIÓN

1.1. Justificación del tema

La Inteligencia Artificial Generativa (IAG) supone uno de los avances más innovadores de los últimos tiempos, dejando paso a una nueva era totalmente desconocida y abierta a descubrir. Destaca por la capacidad de crear contenidos como texto, música, gráficos, vídeo e imágenes estáticas, posicionándola en el ojo público debido a la transformación social y profesional que está ocasionando en la industria audiovisual y del entretenimiento.

En concreto, este proyecto se enmarca dentro del área de la generación de imagen estática para poder evaluar los modelos que producen estas creaciones y que han supuesto una revolución en diversos ámbitos. Por un lado, redefinen el concepto artístico que se conoce hasta la fecha gracias a las nuevas posibilidades creativas y comerciales, permitiendo a artistas y empresas explorar ideas inéditas con una facilidad sin precedentes. Sin embargo, análogamente presentan dilemas morales tales como la autenticidad de las obras generadas, el uso irresponsable de estas herramientas y los derechos de la propiedad intelectual.

En este contexto, es necesario entender el funcionamiento y la estructura que plantean, la motivación que hay detrás de las compañías que han apostado por el desarrollo de estos sistemas, las distintas comunidades que dan soporte y los sustentan y las implicaciones que pueden llegar a tener.

A través del análisis comparativo de los cuatro modelos más importantes a nivel global, es decir, DALL·E, Midjourney, Leonardo AI y Stable Diffusion, y de la investigación de las plataformas que los albergan, se pretende alcanzar todo ello, no solo estudiando las aptitudes y limitaciones técnicas de cada uno, sino también explorando las aplicaciones prácticas y los efectos que causan en la sociedad, con la finalidad de cumplir así con los objetivos propuestos en este TFG.

La temática elegida no es casualidad. Como estudiante del Doble Grado de Ingeniería de Telecomunicaciones y Comunicación Audiovisual, se debe a un interés personal sobre la compenetración tan real que está experimentando la tecnología y la creatividad en este siglo. Ofreciéndonos un instrumento con el que ampliar las habilidades humanas y reflexionar sobre cómo podemos integrarlo en nuestras vidas en lugar de renegar de él, ya que, en este momento, todo apunta a que la Inteligencia Artificial (IA) será la predominante en el sector que nos concierne y la clave para el futuro próximo.

1.2. Objetivos

Evaluar el funcionamiento y sistema de los modelos de Inteligencia Artificial Generativa de imagen estática de DALL·E, Midjourney, Leonardo AI y Stable Diffusion.

Este objetivo general, se concretará a partir de una serie de objetivos secundarios:

1. Investigar la situación actual de la Inteligencia Artificial Generativa en el ámbito social y audiovisual.
2. Diferenciar la Inteligencia Artificial Generativa de otro tipo de IAs, así como definir sus divisiones internas.
3. Exponer la importancia y consecuencias que tiene la generación de imágenes en el panorama tecnológico.
4. Entender los modelos de Inteligencia Artificial Generativa de imagen estática mediante una comparativa entre ellos.

1.3. Metodología

Este proyecto se basa en una metodología dual que combina la investigación bibliográfica y el análisis de interfaz de los modelos de generación de imagen estática que se definen para este informe.

En primer lugar, se efectúa una búsqueda exhaustiva de información para cubrir el marco teórico en el que se encuentra la temática analizada, mediante la lectura de artículos, webs y el visionado de alguna charla de profesionales del sector. Todas ellas, han conformado relevantes fuentes de datos para abordar la cuestión que se plantea en este estudio.

También se ha de tener en cuenta, que se han rechazado muchas de ellas porque trataban aspectos pocos relevantes o centrados en otras áreas enfocadas en características que se alejaban de los objetivos de este TFG. Así pues, esta mecánica ha servido para focalizar las líneas de interés y fijar los límites de este escrito.

En segundo lugar, se especifica el desempeño de los modelos actuales que hay en el mercado y se detallan los cuatro más importantes que se plantean como objeto de estudio para esta memoria:

- DALL·E versión 3, diseñado por OpenAI.
- Midjourney versión 6, diseñado por Midjourney.
- Leonardo.Ai versión Phoenix, diseñado por Leonardo.Ai.
- Stable Difussion versión 3, diseñado por Stability.AI.

Conforme a la documentación recopilada en la primera parte de esta metodología, se procede a desarrollar un análisis individual de cada uno de ellos desde las plataformas de las entidades que los albergan, para luego, diseñar una comparativa en conjunto. Dicho análisis se expresa en 4 fases:

- Acceso al programa: testeo de la interfaz de la aplicación para valorar sus facilidades, conocer su interactividad, funciones y ver si es intuitiva.
- Generación de *prompt*¹: evaluar la precisión y efectividad de los comandos enviados a la IA con el fin de demostrar cómo los modelos pueden crear contenido artístico, lo que favorecerá a su comparativa.
- Resultado: identificación del realismo de las respuestas para su siguiente revisión en base a las funciones proporcionadas por el sistema y determinar así la calidad, los posibles cambios y las limitaciones.
- Corrección: ajustar los parámetros de generación mediante las herramientas que se incluyan para observar las capacidades del modelo.

Para realizar dicha comparativa, se solicitan los siguientes *prompt* en los cuatro casos:

little brunette girl eating a strawberry in a field full of white flowers in the morning, realistic lighting, high quality.

little brunette girl eating a strawberry wearing a blue shirt in a field full of white flowers in the morning, realistic lighting, high quality.

little brunette girl eating a strawberry in a field full of white flowers in the morning, realistic lighting, high quality, 1600s, oil painting –ar 3:2 (este último comando solo para Midjourney).

Los contenidos se producen durante el mes de julio de 2024.

De este modo, el proceso seguido a partir de la experimentación en primera persona y el análisis individualizado de los modelos propuestos permite una percepción más amplia de la generación de imagen estática con Inteligencia Artificial.

¹ Comando que se le da a un sistema de IA para que pueda interactuar y realizar la instrucción solicitada (Fernández, 2023).

1.4. Etapas

A la hora de organizar la propuesta de trabajo, se sigue una estructura que viene determinada por varias etapas que conforman el tratamiento de las sucesivas páginas.

En el marco teórico se desarrolla todo el proceso de documentación sobre la historia de la IAG, cómo nació, de dónde viene, cómo fue creada, al igual que su trayectoria dentro de la sociedad y el mundo audiovisual y las consecuencias que está teniendo en este ámbito.

Asimismo, se explica su programación interna y su método de ejecución. Se diferencian las clases que hay, atendiendo a los distintos algoritmos que implementan y la utilidad que tienen.

Por otro lado, la revolución de la generación de imágenes por Inteligencia Artificial es llevada a cabo por modelos capacitados para idear cualquier concepto que se quiera plasmar, por lo que los más valiosos y potentes pertenecen a empresas tecnológicas multimillonarias que los han dejado a la luz del público, cuestión que también se aborda aquí.

Una vez conformado esto, se dispone a confeccionar un análisis particular de cada modelo establecido (DALL·E, Midjourney, Leonardo AI y Stable Diffusion), en el cual se valora la plataforma en la que se encuentra, se examina su interfaz, las funciones que poseen, la complejidad de los *prompt*, el nivel de respuesta, la comunidad en la que se basan y la calidad del resultado, siguiendo así el planteamiento propuesto en la metodología.

Con todo ello, se realiza una comparativa donde se podrán observar las ventajas y desventajas, las similitudes y diferencias y así apreciar el enfoque y estrategia que ha suscitado cada compañía.

Finalmente, se deducen las posibles conclusiones con un criterio objetivo y se exhiben los recursos bibliográficos consultados.

2. LA IA GENERATIVA EN EL AUDIOVISUAL

2.1. Primeros años

El concepto de Inteligencia Artificial aparece por primera vez en 1950 de la mano de Alan Turing, considerado el padre de la ciencia de la computación, al publicar

su artículo académico *Maquinaria computacional e inteligencia*² (McCarthy, 2007)

Seis años después, el informático e investigador John McCarthy acuña el término en la primera conferencia que se realizó sobre IA en el Dartmouth College (New Hampshire, Estados Unidos).

En 2007, cuatro años antes de morir, dejó escrito en uno de sus últimos artículos para la Universidad de Stanford una definición sobre esta disciplina que todavía perdura hasta la fecha:

Es la ciencia y la ingeniería para crear máquinas inteligentes, especialmente programas informáticos inteligentes. Está relacionada con la tarea similar de utilizar ordenadores para comprender la inteligencia humana, pero la IA no tiene por qué limitarse a métodos que sean biológicamente observables (McCarthy, 2007, p. 2).

De manera que, conforme el tiempo ha ido pasando, la sociedad ha ido evolucionando y con ella la tecnología. A partir de 2012, las grandes multinacionales tecnológicas deciden apostar por el desarrollo de la IA, invirtiendo en el aprendizaje profundo después de la publicación en la Universidad de Toronto de AlexNet, un modelo de red neuronal creado por Geoffrey Hinton, Alex Krizhevsky e Ilya Sutskever que podía clasificar 1,2 millones de imágenes de alta resolución en 1.000 clases diferentes. Este proyecto ganó el *ImageNet Large Scale Visual Recognition Challenge*, un desafío que se celebraba anualmente para conseguir diseñar softwares de reconocimiento visual de objetos con el mínimo error posible. En concreto, AlexNet presentó una tasa de error del 9,8% más baja por encima de la del subcampeón (Briggs & Carnevali, s. f.).

Este modelo no hubiera existido si no hubiera sido por la disponibilidad de grandes conjuntos de datos que ya almacenaba ImageNet y por el avance en la aceleración de procesamiento gráfico mediante GPUs (Rosso, 2019).

No es hasta que ocurre un hito de este tipo cuando el mundo empieza a darse cuenta del potencial que tienen ciertas técnicas y procesos que estaban ahí pero no se les había prestado atención.

Así pues, desde aquí en adelante, la IA empezó a diversificarse en muchas ramas y se sumergió completamente en la esfera del audiovisual, ya que muchos sectores de este la empezaron a demandar firmemente porque agilizaba bastante el trabajo y los resultados eran prácticamente instantáneos (Rosso, 2019).

² En él, se aborda la cuestión de si las máquinas pueden pensar y se propone una forma de evaluar la inteligencia de una computadora mediante lo que hoy se conoce como el test de Turing.

2.2. Contexto histórico y social

La Inteligencia Artificial Generativa (IAG) se debe diferenciar de cualquier otro subcampo de IA porque es la única capaz de crear contenido nuevo y original, ya sea escrito, visual o auditivo. Se aleja así de la IA analítica, tal vez la mayor establecida hasta la fecha, que se encarga principalmente de reconocer y clasificar materia existente. Las capacidades cognitivas se transforman así en capacidades creativas (Barba & Valenzuela, 2023).

Nace como extensión de los modelos *Machine Learning* y *Deep Learning*, al analizar información existente de la red para poder generar el producto que se desea o simplemente elaborar nuevas muestras en función de otras que poseen características parecidas y se encuentran en el conjunto de datos con el que fueron entrenados (Hidalgo, 2024).

A mitad de la década de los 2010, después del *boom* que la IA había alcanzado hasta el momento y con la configuración que se había desarrollado para poder hacer frente a esta tecnología, Ian Goodfellow introdujo las Redes Generativas Antagónicas (GANs), que permitieron producir las primeras imágenes realistas (Lawlor & Chang, 2024).

Al poco tiempo, Google compartió con el mundo el Transformer, una nueva arquitectura que revolucionó el procesamiento del lenguaje natural y sirvió para sentar las bases de modelos más avanzados.

En menos de dos años, entre 2018 y 2020, la multinacional Open AI lanzó un modelo de texto (y su posterior versión mejorada) basado en estos transformadores conocido como ChatGPT, el cual consiguió realizar escritos tan bien cohesionados y precisos que dificultó distinguirlos de los desempeñados por humanos (Lawlor & Chang, 2024).

La empresa pudo trabajar a fondo y definir detalle a detalle todas sus estructuras durante la pandemia COVID-19, cuando la paralización de la mayoría de los servicios jugaron a su favor. De manera que, cuando se recuperó la normalidad, este tuvo un impacto tan grande en la sociedad que a día de hoy es el producto más utilizado en cualquier compañía, por lo que su expansión y uso comercial sigue en auge actualmente, incorporando cada vez nuevas funciones como la creación de imágenes (DeVon, 2023).

A raíz de esto, la competencia quiso alcanzar el mismo número de usuarios que estaba moviendo Open AI, así que en el tercer sector se inició una evolución en poco tiempo innovando en más herramientas no solo de generación de imágenes, sino de audio, de video, de diseño gráfico, efectos visuales y entre otros (Singla et al., 2023).

En 2023 existían el triple de modelos de IAG que en 2020/21, debido a que las empresas MAGMA (Microsoft, Apple, Google, Meta y Amazon) empezaron a invertir en ellos (Singla et al., 2023).

Todo este rápido crecimiento dio lugar a muchas dudas dentro de la población y sobre todo en la industria creativa profesional.

Una vez más, la tecnología avanza a una velocidad tan rápida que las leyes de los gobiernos no pueden seguirle el ritmo y surgen problemáticas al respecto.

En el ámbito de la IAG, de acuerdo con el medio Página12 (2023), diversos artistas se han manifestado en varios medios y redes sociales sobre lo que supone que un algoritmo entrenado para imitar un estilo artístico reproduzca en cuestión de segundos un concepto visual que puede tardar varios meses en completarse a mano. Como es el caso de la dibujante estadounidense Sarah Andersen, quien fue una de las primeras personas en llevar esta disyuntiva frente a los tribunales federales en defensa de los derechos de autor.

Gracias a ella, toda una generación de creadores de contenido visual salió en su apoyo, algunos se sumaron a su demanda colectiva y otros empezaron a denunciar situaciones profesionales de la misma índole que estaban sufriendo. Se formaron miles de plataformas y grupos alrededor de todo el planeta para abogar por la misma causa.

La finalidad que ellos buscan no es rechazar la IAG, ya que esta ha venido para quedarse y es prácticamente imposible eliminarla, sino coexistir con ella. Por ese motivo, piden a las grandes compañías que manejan estos modelos que se les dé una compensación cuando alguien genere una imagen con su estilo y que, además, puedan rechazarla o aceptarla. No obstante, este dilema aún sigue abierto sin resolverse (Página12, 2023).

Por otra parte, poco después, cantantes y compositores musicales se vieron involucrados en el mismo contexto. Este mismo año, más de 200 artistas reconocidos y emergentes han firmado una carta, alentada por la organización Artist Right Alliance³, donde se expresa la preocupación sobre cómo se van a tratar los derechos creativos de sus obras frente al uso inadecuado de herramientas que crean temas con sus voces, letras y melodías (Palomino, 2024).

Algunas productoras ya han empezado a tomar medidas frente a ello, cuando se han dado cuenta que podían perder grandes beneficios y las propiedades intelectuales que poseen verse perjudicadas.

El problema es que tienen que buscar soluciones más rápido que las propuestas o casi inexistentes que plantea la Administración Pública. Así que, la única

³ Asociación estadounidense sin ánimo de lucro que lucha por los derechos de los creadores de la industria musical en el sector digital (Palomino, 2024).

manera de realizar reivindicaciones es agruparse para formar la mayor fuerza posible y conseguir que se resuelva con unas condiciones adecuadas (Palomino, 2024).

Un claro ejemplo, es la lucha del Sindicato de Guionistas de Estados Unidos (WAG) que, junto al de actores (SAG-AFTRA), iniciaron una manifestación a mediados de 2023 que duró más de 4 meses y paralizaron toda la industria de Hollywood. En ella se pedían mejores condiciones laborales y sobre todo una mejor regulación de la IA. No querían que guiones e historias fueran escritas por herramientas inteligentes y que las productoras se pudieran ahorrar horas de trabajo y grandes costes en contratación de personal. Tampoco querían que imágenes y vídeos de actores pudieran ser replicados en películas y series sin que esa persona lo permitiera u obtuviera un beneficio. Por lo que, tras semanas dejando sus ocupaciones para ir a alzar sus voces en las calles de Los Ángeles y Nueva York, finalmente consiguieron firmar con la Alianza de Productores (AMPTP) un acuerdo que cumplía todos esos requisitos (García, 2023).

A pesar de que la industria perdió 6.500 millones de dólares y costó 45.000 puestos de trabajo (Ulgu-Servant, 2023), se puede decir que todas esas personas fueron visionarias y adelantadas a su tiempo, ya que abordaron una gran problemática cuando nadie lo iba a hacer por ellos.

Por su parte, la Unión Europea ha aprobado la primera ley de IA de la historia, la cual entrará en vigor en 2026. Con ella se pretende tener una mejor regulación de la tecnología en el mercado laboral europeo y fomentar un entorno digital seguro. Así lo afirma el eurodiputado italiano Brando Benifei (2024):

Finalmente tenemos la primera ley vinculante del mundo sobre Inteligencia Artificial, para reducir riesgos, crear oportunidades, combatir la discriminación y aportar transparencia. Gracias al Parlamento, se prohibirán en Europa prácticas inaceptables de IA y se protegerán los derechos de los trabajadores y los ciudadanos.

Se espera que los estados miembros cumplan con la normativa anticipándose de este modo a otras principales potencias como Estados Unidos o China. La vulneración de algunos de los puntos del nuevo reglamento supondrá multas que pueden alcanzar hasta los 35 millones de euros (Alonso, 2024).

2.3. Estructura y funcionamiento

Para comprender mejor como actúa la IAG, se debe entender su funcionamiento a través de su implementación y arquitectura.

Existen dos modelos base de los que parte cualquier tipo de IA: el *Machine Learning* y el *Deep Learning*.

Respecto al primero, se trata del aprendizaje automático, es decir, la facultad que tienen las computadoras de poder pensar por sí mismas mediante el desarrollo de algoritmos y procesos que les permite aprender a partir de datos. Debido a esto, no se necesita de la supervisión humana porque las máquinas son capaces de relacionar e identificar patrones que les posibilitan tomar ciertas decisiones y realizar predicciones (Rodríguez, 2020).

Estos algoritmos son entrenados para ser perfeccionados con el análisis de una gran cantidad de datos en la red. La finalidad es conseguir la mayor precisión y los mejores resultados posibles.

Por otro lado, el *Deep Learning* es una subcategoría del *Machine Learning*. Se inspira en la estructura y la actividad del cerebro humano, donde trabaja con redes neuronales profundas que se clasifican en diversas capas y cada una de ellas realiza una tarea de análisis concreta. Al procesar la información correspondiente, pasa al siguiente nivel hasta que se obtiene la solución requerida. Todas las capas están conectadas entre sí y fluyen en la misma dirección, como se puede ver en la Figura 1.

Los algoritmos de estas redes también se someten a entrenamientos con enormes cantidades de datos que les ofrecen poder reconocer patrones de datos no etiquetados, que a diferencia del *Machine Learning*, sí que necesita categorizar previamente de qué tipo es cada elemento que recibe.

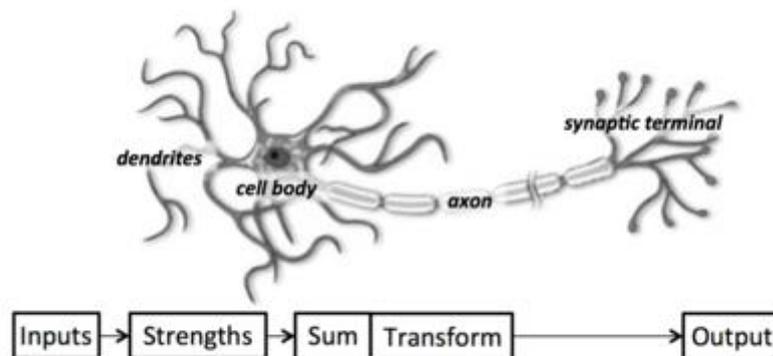


Figura 1. Estructura de una red neuronal profunda en *Deep Learning*. Extraído de Xataka. <https://www.xataka.com/robotica-e-ia/machine-learning-y-deep-learning-como-entender-las-claves-del-presente-y-futuro-de-la-inteligencia-artificial>

Actualmente, empresas como Tesla y Uber están haciendo uso de este modelo en su implementación de la conducción autónoma en algunos de sus vehículos.

A través de la identificación de variables del exterior en las carreteras y del comportamiento humano delante del volante, se pretende reproducir un patrón de conducción tan exacto al de una persona real que garantice que este hecho

se pueda dar al 99,999% de precisión sin ningún error. Incluso predecir accidentes antes de que sucedan.

Este es solo un caso de cómo se está empleando el *Deep Learning*. Es el modelo preferido en miles de compañías para trabajar con IA, pero aún le queda una larga evolución y un gran reto por delante (Rodríguez, 2020).

Luego de su desarrollo, surgieron otros dos modelos innovadores de esta rama que dieron pie al nacimiento de la IAG en 2014: las Redes Generativas Antagónicas (GANs) y los Transformers (Lawlor & Chang, 2024).

Ambas parten de la arquitectura del *Deep Learning*. Las GANs se sirven del aprendizaje profundo por medio de dos redes neuronales que compiten entre sí. Mientras que una toma el papel de red generadora, produciendo nuevos datos a partir de un conjunto de otros ya entrenado, la otra se convierte en la red discriminadora, la cual trata de averiguar si la salida de datos resultante es falsa o verdadera, en definitiva, si los datos generados pertenecen a ese conjunto entrenado o no (Figura 2) (Amazon Web Services [AWS], 2023).

Las dos se necesitan mutuamente para poder formarse como GAN, ya que, con cada iteración, cada componente mejora su capacidad de acuerdo con el rendimiento del otro.

Se considera que una GAN es exitosa cuando el proceso de formación llega a su fin, en el cual el discriminador ya no reconoce si los datos sintetizados por el generador son reales o artificiales.

Se consigue un estado de equilibrio interno cuya técnica ha servido para elaborar contenido genuino como texto, imágenes, vídeos, gráficos o sonidos (AWS, 2023).

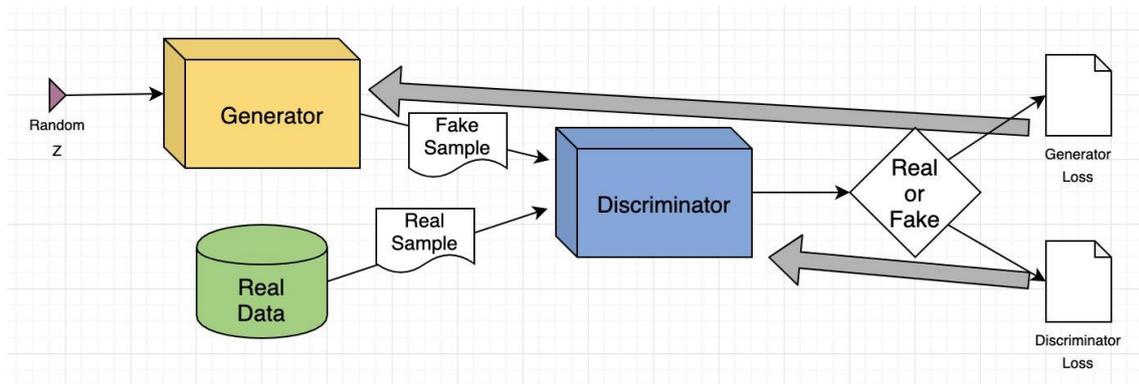


Figura 2. Arquitectura interna de una GAN. Extraído de AWS. <https://aws.amazon.com/es/what-is/gan/>

En cuanto a los Transformers, son también redes neuronales que aprenden el contexto y las relaciones de las palabras de los *prompt* de entrada (Bengesi et al., 2024).

Con estas peticiones se establece una comunicación usuario-máquina con la que se lleva a cabo el resultado generado (Bengesi et al., 2024).

Cuanto más definido sea el *prompt*, más detallado será el producto final. Es por eso que hay páginas web que venden este tipo de información porque a veces desbloquean funcionalidades que no se sabía que esa IA podía alcanzar (Fernández, 2023).

Por medio de algoritmos matemáticos, los transformadores identifican la importancia y la conexión de los términos para crear la secuencia de salida. Son la pieza primordial para interactuar con una IA (Bengesi et al., 2024).

Gracias a ellos, la IAG se ha podido desenvolver a gran velocidad con estos modelos, superando y reemplazando en muchas ocasiones a estructuras *Deep Learning* creadas hace cinco años (Merritt, 2022).

Asimismo, no solo han cambiado el panorama del procesamiento del lenguaje natural con los mecanismos pioneros del *Machine Learning*, sino que han ayudado a lograr el incremento de la calidad y la variedad de lo que se puede idear (Merritt, 2022).

2.4. La generación de imágenes

Con la llegada de las GANs, el concepto de Inteligencia Artificial Generativa se desveló para el mundo entero.

Como su nombre indica, puede originar cualquier tipo de contenido visual o auditivo, vía unos algoritmos que sintetizan las palabras que el usuario les proporciona y las transforman en un producto muy exacto y preciso (Barba & Valenzuela, 2023).

Entre ellos se encuentran videos, audios, efectos, textos e imágenes.

Bajo la implementación de esta nueva arquitectura, las empresas empezaron a apostar por modelos que utilizaban estas redes, las cuáles evalúan cómo de verosímiles son los datos para poder generar las ilustraciones (MIT Technology Review, 2023).

Los resultados llegaron a ser tan realistas y detallados que miles de personas querían tener esa tecnología incorporada en sus vidas inmediatamente (MIT Technology Review, 2023).

Las compañías comenzaron a ganar grandes beneficios por el uso de sus modelos, que además no tenían ningún tipo de restricción. Cualquier internauta podía crear lo que quisiera (Rico, 2023).

Los gobiernos tampoco tenían ninguna ley ni decreto que pudiera regular este tipo de IA y, al poco tiempo, las redes sociales y un sinfín de plataformas se vieron inundadas con imágenes desarrolladas artificialmente.

Entonces es cuando se manifestó la gran problemática que acarrea este tema.

Es muy fácil poder escribir lo que uno desea en una caja de comentarios y al cabo de unos segundos obtener un retrato con un estilo idéntico al de Da Vinci o Picasso, por ejemplo.

Debido a esto, el panorama del sector artístico se modificó por completo, las creatividades tomaron un nuevo enfoque, sobre todo en el entorno digital, donde se han vendido obras de todas las clases, incluso dentro de la moderna tecnología de las *blockchain*⁴ en formato NFT⁵ (Sainza, 2023).

En cambio, esto perjudicó a muchos artistas que, como ya se ha expuesto anteriormente, vieron sus trabajos plagiados con imágenes que recreaban sus estilos personales y que les había costado horas y meses de elaboración.

Surgieron transformaciones también en los campos de la publicidad y el marketing con anuncios personalizados en el diseño gráfico y en el ámbito del cine y los videojuegos con creaciones de escenarios, entornos y objetos, entre otros (Rico, 2023).

Desde ArtStation, una de las mayores comunidades de diseñadores visuales que existe a nivel global, ya han denunciado la mala práctica que se tiene de la IAG, porque no es cuestión de que quite puestos de empleo por la automatización que se puede llegar a tener con esta, sino porque se basa en datos que supuestamente están protegidos con derechos de autor para generar el producto final, sin tan siquiera reconocer la fuente original (Díaz, 2024).

Así pues, en la mayoría de los países, debido a las leyes que resguardan los derechos de los trabajadores, no pueden perder sus ocupaciones por la entrada inmediata de la IA, la cual era una de las cuestiones más temidas a nivel social.

Por lo que, aunque en un principio se pensaba que las empresas buscaban echarlos, en verdad sucedía todo lo contrario, quieren personas que sepan y entiendan el funcionamiento de estas innovadoras técnicas para incrementar las

⁴ Tecnología de registro distribuido que posibilita preservar un historial de transacciones de forma, transparente, segura e invariable (Pereyra, 2024).

⁵ Del inglés Non-Fungible token, se trata de una clase de información encriptada que no puede ser suplantada por otra ya que tiene un distintivo único (Cabello, 2021).

capacidades humanas y de la compañía, de acuerdo con Barba y Valenzuela (2023).

Adicionalmente, cabe recalcar otro punto que ha causado complejidades en la sociedad, los *deepfakes* o imágenes/videos falsos.

Estos hacen creer que alguien relata un suceso cuando en verdad no lo ha hecho. Utilizan su cara y su voz a partir de un entrenamiento basado en data que se ha recopilado de ese individuo para modificarlo y construir la narrativa deseada. Lo mismo pasa al generar imágenes de personas realizando una acción, portando algo o estando en un lugar distinto al real (National Cybersecurity Alliance, 2023).

Celebridades, políticos, deportistas y otras figuras públicas han sido víctimas de ello (Figuras 3 y 4).



Figura 3. *Deepfake* del Papa Francisco con un abrigo de la firma Balenciaga. Extraído de Pablo Xavier, La Vanguardia. <https://www.lavanguardia.com/tecnologia/20230328/8859266/creador-imagen-viral-papa-francisco-chaqueta-deep-fake-pmv.html>



Figura 4. *Deepfake* de los expresidentes de los Estados Unidos, Barack Obama y Donald Trump, jugando un partido de baloncesto. Extraído de MDTECH. <https://www.mdtech.news/servicios/2023/3/24/esta-ia-recreo-donald-trump-barack-obama-jugando-al-baloncesto-8456.html>

La única forma de contrarrestar este problema es programando un modelo que pueda identificar las falsedades del contenido y denunciarlas. Sin embargo, hasta llegar ahí, lo único que se puede hacer a nivel individual es navegar de forma segura con los ajustes de privacidad activados, evitar cualquier intento de *phishing* y valorar de forma crítica todo lo que se vea (National Cybersecurity Alliance, 2023).

Poco a poco, los gobiernos están trabajando en estas regulaciones y llegará un momento en el que se controle gran parte del desempeño de la IA, ya que, sin los humanos, esta no puede funcionar y las empresas disminuyen sus ganancias (Sainza, 2023).

2.5. Modelos actuales

Existe una amplia variedad de modelos de generación de imagen estática que han tenido una gran evolución para la poca trayectoria que han recorrido en los últimos años.

Los estudiados aquí se han convertido en los más relevantes debido al elevado nivel de rigor y la considerable precisión a la hora de reproducir las ilustraciones tan detalladas y realistas (MIT Technology Review, 2023).

Han demostrado ser una innovación para la tecnología que conforma el siglo XXI y un referente del claro ejemplo que supone la IAG en el sector audiovisual.

Según estiman Chui et al. (2023), esta añadirá a la economía global entre \$2.6 y \$4.4 billones en valor anual, ampliando el impacto económico del conjunto de la IA entre un 15 y un 40%.

En consecuencia, se podría afirmar que las multinacionales que disponen y construyen estos proyectos, son las que poseen las llaves del futuro.

Concretamente, los modelos que se exponen son, en este orden: DALL·E, Midjourney, Leonardo AI y Stable Difussion.

DALL·E

Recibe su nombre en honor al pintor español Salvador Dalí por concebir imágenes artísticas y únicas y al personaje de Disney WALL·E (Lin, 2023).

Pertenece, junto con ChatGPT y Sora, a los modelos de IAG desarrollados por OpenAI, la mayor empresa especializada en LLMs⁶ y fundada en 2015 por dos de los multimillonarios americanos más destacados del planeta, Elon Musk y Sam Altman.

DALL·E parte de una arquitectura de Transformer, donde una entrada de texto genera un icono pictórico con las especificaciones que incluyen esos parámetros.

Este tipo de lenguaje garantiza la relación entre los conceptos que se le otorgan para llevar a cabo la respuesta (Lin, 2023).

Su primera versión fue lanzada en 2021, su segunda en 2022 y su tercera y más reciente en 2023. En ella se puede encontrar una mejor precisión en los *prompt* al manejar la herramienta, ya que se ajusta mejor a las descripciones textuales sin la necesidad de escribir tantas palabras a la hora de solicitar lo que se busca, y sin tampoco tener que recurrir a la ingeniería de *prompt* para conocer los términos exactos que la máquina usa para conectar los significados que le ofrece el usuario con los que ella utiliza para describir cierto aspecto (Fernández, 2024).

Además, se han actualizado algunos detalles que pecaban de errores como las manos y dedos de las personas generadas.

También puede crear imágenes a partir de otras imágenes, otorgándoles estilos y ratios distintos conforme se quiera (Fernández, 2024).

Se alberga en la plataforma de OpenAI como extensión al modelo de texto ChatGPT y tiene un coste de 20€ al mes (OpenAI, 2024).

⁶ Por sus siglas en inglés, *Large Language Models*, hacen referencia a modelos que pueden realizar inmensas tareas a la vez debido a su entrenamiento en volúmenes enormes de data (International Business Machines [IBM], 2024).

La versión 3 se encuentra disponible en fase beta para las cuentas de pago. No obstante, se puede hacer un uso gratuito de ella a través de Bing Chat, la IA de Microsoft (Fernández, 2024).

MIDJOURNEY

Es quizá el competidor más potente en el mercado de la IAG por sus imágenes hiperrealistas y la claridad de las definiciones en su base de datos. Creadores de contenido, marcas y productoras recurren a él diariamente para hallar el mejor resultado posible (Himel, 2024).

Nace de la mano de David Holz, co-fundador de la *start-up* de realidad virtual y aumentada, LeapMotion. A este se le fueron sumando otros prestigiosos intelectuales del mundo de la programación que habían trabajado para Tesla, Intel y Apple, entre otros (Dircomfidencial, 2023).

De este modo, formaron un laboratorio de investigación de nombre homónimo en el cual se dedican a analizar el comportamiento de la IA y su perfeccionamiento para un avance global.

El sistema estuvo en fase cerrada durante varios meses antes de abrirse al mundo en 2022 (Dircomfidencial, 2023).

En el poco período que lleva público, ha conseguido ser de los más rentables, según afirman Wiggers y Coldewey (2024), para TechCrunch, logrando recaudar 200 millones de dólares como beneficios sin ninguna inversión proveniente del exterior.

Emplea la misma técnica que el modelo anterior, el *text to image*, que transforma una línea de palabras relacionadas en una ilustración (Dircomfidencial, 2023).

Cabe señalar, las numerosas versiones que han ido saliendo conforme la plataforma iba profundizando en sus puntos fuertes. En el año de debut se estrenaron las primeras cuatro versiones en cuestión de seis meses. Cada una superando a la anterior hasta llegar a 2024, donde se tiene la v6 alfa, la cual se ha viralizado en la industria del entretenimiento por su alta exactitud, la corrección de errores y la capacidad de originar a una resolución de 1792 x 1024 píxeles (Fernández, 2024).

Midjourney se puede localizar mediante un bot en la app de Discord y tiene una tarifa mensual de 10, 30, 60 o 120€ y anual de 96, 288, 576 o 1152€, dependiendo de sus funcionalidades (Midjourney, 2024).

LEONARDO AI

Modelo al cual le da nombre el célebre pintor e inventor del Renacimiento italiano Leonardo Da Vinci.

Surgió con el objetivo de democratizar el arte y empoderar a los creadores con la última tecnología de vanguardia. A través de una interfaz sencilla, se pretende conseguir que cualquier usuario, independientemente del nivel de conocimiento que disponga sobre IA, pueda generar cualquier resultado (Raza, 2024).

Sus fundadores, JJ Fiasson, Chris Gillis y Jachin Bhasme, tres eruditos en las finanzas, decidieron erigir una compañía australiana que trabajara con contenido de activos para videojuegos con ayuda de inteligencia artificial, pero rápidamente se dieron cuenta del impacto que estaba teniendo la IAG en el mundo y su idea tomó otro camino (Shu, 2023).

En 2022, Leonardo AI salió al mercado apostando por una estrategia vertical y el modelo de negocio B2B⁷.

Con la ayuda de inversionistas como AWS o Blackbird Ventures, ha crecido estos dos últimos años consiguiendo facturar una cifra de 30,7 millones de dólares (PitchBook, 2024).

Por otro lado, el modelo de trabajo sigue siendo igual al descrito en los otros dos casos, en el que, partiendo de unos términos, se construye una imagen. Prácticamente todos funcionan con la misma mecánica.

Aun así, hay que tener en cuenta que este aprovecha varios motores de entrenamiento que proceden de otras desarrolladoras, como Pixel Art, Vintage o incluso dos de Stable Diffusion (El Español, 2023).

Los internautas pueden escoger en base al estilo que quieran realizar el que mejor les convenga, sin el requisito de escribir comandos complejos para hacerse entender con el sistema (El Español, 2023).

Asimismo, se subraya la implementación de *Live Canvas*, una función que permite dibujar un boceto de la figura cuando el usuario introduce un primer mensaje y que, a partir de este, se pueden ir añadiendo, eliminando o cambiando elementos en tiempo real (Shu, 2023).

Igualmente, la integración de API (Interfaz de Programación de Aplicaciones) involucra a otras profesiones como desarrolladores, al facilitarles la incorporación de los servicios de IA de este modelo en otros sitios webs y apps. El código se encuentra en la página del programa (Huang, 2024). Lo mismo

⁷ *Business tu business* o “negocio a negocio” es una clase de funcionalidad entre empresas, en las que las transacciones comerciales se establecen entre ellas (Sánchez, 2020).

ocurre con DALL·E. Sin embargo, Midjourney no cuenta con una API oficial, sino que está constituida por terceros.

Leonardo AI se aloja en la plataforma homónima con su versión gratuita de 150 tokens⁸ al día. Cada uno concede la generación de una imagen, aunque también se pueden adquirir más por un precio de 10, 24 o 48€ (Leonardo.AI, 2024).

STABLE DIFUSSION

Se distingue de todos los demás por ser un modelo de código abierto, esto quiere decir que cualquier programador tiene la oportunidad de construir herramientas y funciones partiendo de él, por lo que, posibilita tener una comunidad bastante amplia para todos los usuarios en la que se puede trabajar conjuntamente para mejorar el desarrollo de esta IA (Lara, 2023).

De la misma manera que Leonardo AI, la filosofía de la empresa que hay detrás aboga por la democratización de las nuevas tecnologías. Es por eso que, su fundador, Emad Mostaque, creó la compañía inglesa Stability AI en 2022 con el fin de cumplir este propósito (Alcaraz, 2023).

El motor de renderizado recurre al método de cualquier IA generativa de imagen estática, el *text to image*, que ya se ha explicado con anterioridad, pero además dispone del *image to image*, con el que se puede generar una figura pictórica a partir de otra (Fernández, 2023).

Para servirse de estas dos formas, Yi et al. (2024) afirman que se emplea un proceso distinto a los demás modelos para poder conseguir ilustraciones más realistas y auténticas. Este se divide en dos etapas y se conoce como *Denoising Score Matching* (DSM), una técnica de difusión de ruido basada en los datos de entrenamiento que utiliza para la creación de los resultados.

En primer lugar, se encuentra la fase de adición de ruido directo, en la que se aplica un proceso paulatino de granulado en una imagen real con el fin de deteriorarla hasta convertirla en una distribución de ruido invisible, explícitamente, en una figura indistinguible.

Tras esto, se pasa a la segunda fase, el *denoising*. Aquí se elimina de manera reversible el ruido Gaussiano inyectado en la etapa inicial, a partir de un proceso inverso e iterativo para alcanzar una nueva imagen reconstruida que se asimile a la nativa.

Por ende, gracias al adiestramiento en este procedimiento, el modelo se instruye para saber identificar distorsiones y variaciones en sus solicitudes y proyectar una mejor respuesta (Yi et al., 2024).

⁸ Crédito digital.

Por otra parte, existen estructuras de redes neuronales adicionales, como ControlNet, que se incluyen dentro de Stable Diffusion para enriquecerlo con otras funciones y dotarlo con controles de acondicionamiento espacial, como, por ejemplo, mapas de segmentos, calibradores de bordes, ajustes de profundidad, etc. Así pues, el producto final consigue ser más robusto y equilibrado (Zhang et al, 2023). Se puede acceder a este plugin⁹ desde la página web que comparte el mismo nombre.

El objetivo principal es integrar el *prompt* textual con una imagen prototipo que se sube previamente a la aplicación para extraer las características más representativas y formar una función combinable que proporcione un efecto visual más personalizado (Polo et al., 2023).

Por último, Stable Diffusion se encuentra en su tercera versión y se puede abrir mediante código local o a través de la plataforma gratuita Dream Studio (25 tokens al día). Cuenta con más de 10 millones de usuarios y brinda la opción de incorporar API (Stability.ai, 2024).

3. ANÁLISIS DE LOS MODELOS DE IA GENERATIVA

3.1. DALL·E

- Acceso al programa:

Desde la corporación de OpenAI, se ofrece la posibilidad de ingresar a DALL·E por vía del multimodelo conocido como ChatGPT, el cual alberga esta IA de generación de imágenes.

Para ello, desde la página web (Figura 5), registrándose con un usuario y contraseña, se tiene acceso a dicha herramienta que se sitúa en el menú superior de la izquierda, junto con otro tipo de aplicaciones (Figura 6).

⁹ Función que sirve como extensión de un programa sin la necesidad de modificar el código (Polo et al., 2023).

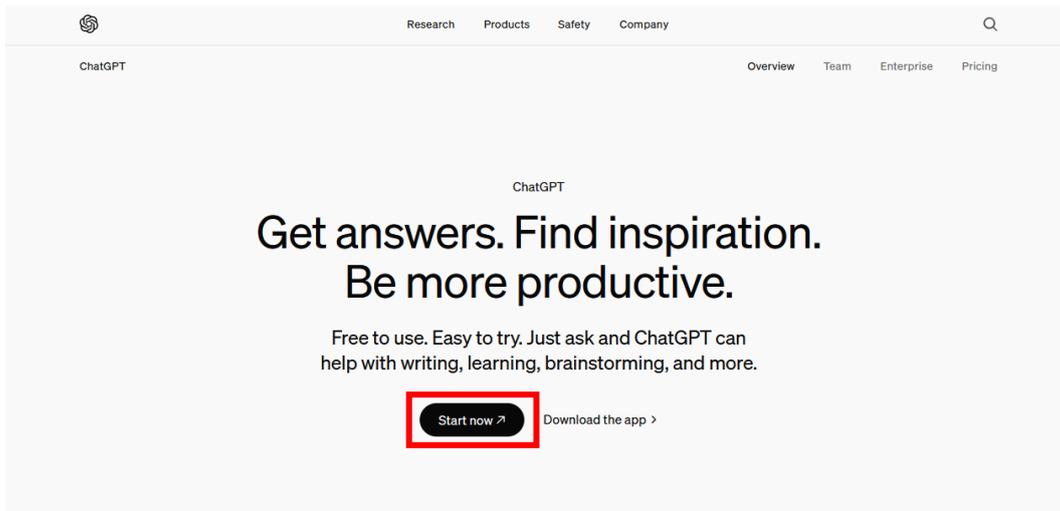


Figura 5. Página web principal de ChatGPT donde se inicia sesión. Extraído de OpenAI. <https://openai.com/chatgpt/>

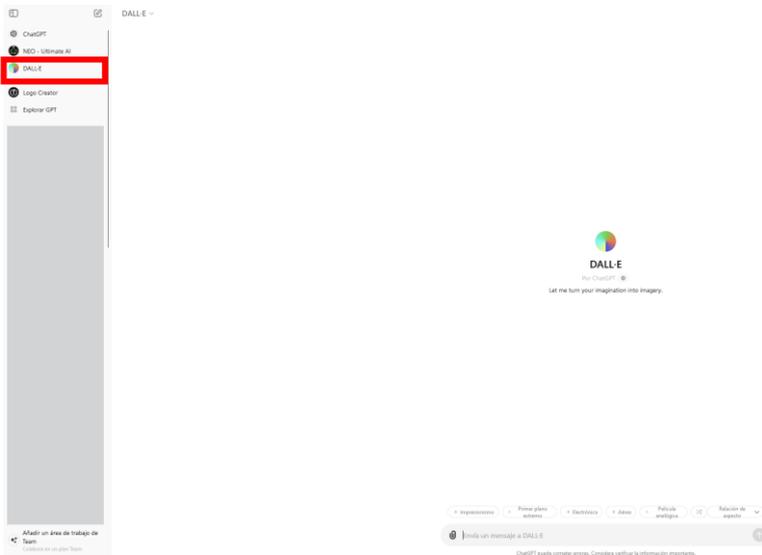


Figura 6. Herramienta DALL·E preparada para su ejecución. Extraído de OpenAI. <https://openai.com/chatgpt/>

- Generación de *prompt*:

Una vez dentro, en la barra de texto de la parte inferior, se introducen las indicaciones necesarias para crear la ilustración deseada. Han de ser claras y concisas para que el sistema reconozca estos términos y los relacione con la mayor precisión posible. De esta forma, el significado visual será más exacto a lo esperado.

Se deben especificar siempre todos los parámetros que vayan a salir en la imagen, es decir, desde el sujeto hasta la localización, los colores, texturas, expresiones, detalles de la iluminación y otros objetos que se requieran.

Los comandos se pueden escribir en cualquier idioma, pero se ha de tener en cuenta que el entrenamiento de estas máquinas se ha realizado en inglés, por lo que, si se aplican en esta lengua, se garantizan mejores resultados.

En este caso, se va a generar una niña pequeña de cabello oscuro comiendo una fresa en un campo lleno de flores blancas a la luz del día (Figura 7). Por lo tanto, el *prompt* que se va a utilizar es: *little brunette girl eating a strawberry in a field full of white flowers in the morning, realistic lighting, high quality.*

- Resultado:



Figura 7. Resultados de imagen estática generados por IA en DALL·E mediante un *prompt* de texto. Elaboración propia a partir de OpenAI. <https://openai.com/chatgpt/>

Por defecto, DALL·E entrega al usuario dos conceptos similares con una proporción de 1:1.

Si la imagen no cumple con las expectativas que se tenía, se puede modificar ajustando o añadiendo palabras al *prompt* o clicando sobre una de ellas y editándola.

A continuación, se prueban ambas posibilidades para ver sus efectos, ahora con la niña vistiendo una camiseta azul (Figuras 8 y 9).

Respecto a la primera opción, no hace falta volver a redactar todo el comando porque el sistema retiene la información dada anteriormente, simplemente se le agrega la nueva solicitud: *she is wearing a blue shirt.*



Figura 8. Resultados de imagen estática generados por IA en DALL·E modificando el *prompt* original desde el cuadro de texto. Elaboración propia a partir de OpenAI. <https://openai.com/chatgpt/>

En la segunda opción, se escoge una ilustración de una de las originales y sobre ella, se introduce la instrucción.

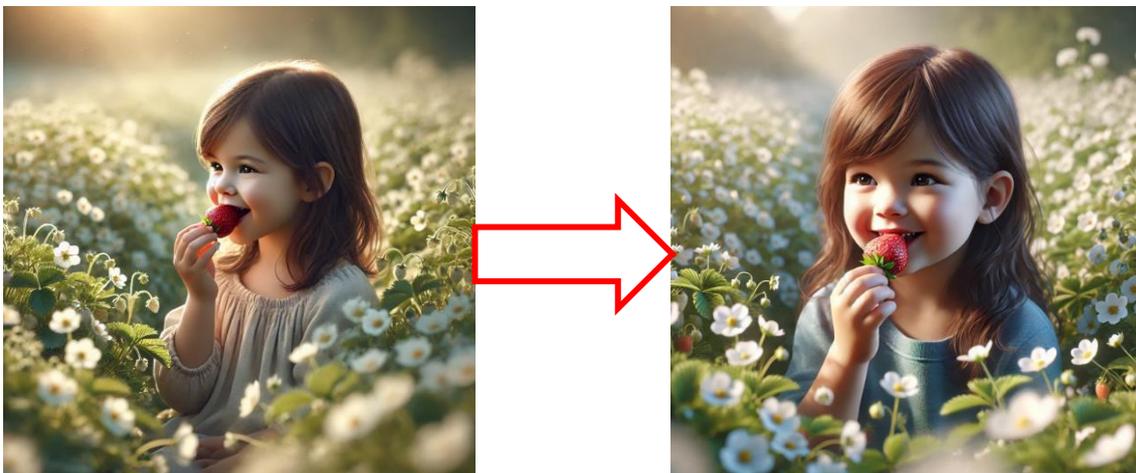


Figura 9. Resultados de imagen estática generados por IA en DALL·E modificando el *prompt* desde la imagen original. Elaboración propia a partir de OpenAI. <https://openai.com/chatgpt/>

En ambos resultados, se cumple con lo señalado, pero el sujeto cambia de enfoque, el tipo de perspectiva no es la misma y algunos rasgos faciales se pierden. El retrato no se mantiene al 100%, sin embargo, la calidad sí.

- Corrección:

Por otro lado, se puede hacer uso directamente de algunas de las funciones predeterminadas que se sitúan abajo (Figura 10), como, por ejemplo, establecer otro ratio o estilo artístico (Figura 11).



Figura 10. Funciones predeterminadas para la modificación de imágenes en DALL·E. Extraído de OpenAI. <https://openai.com/chatgpt/>



Figura 11. Resultados de imagen estática generados por IA en DALL·E modificando el *prompt* original ahora con un ratio 16:9 y un estilo abstracto. Elaboración propia a partir de OpenAI. <https://openai.com/chatgpt/>

Además, cada imagen viene conferida con tres opciones más con las que se pueden interactuar: seleccionar una parte para cambiarla, descargar el archivo y visualizar la solicitud recibida por la IA en formato extenso (Figura 12).

Una gran ventaja de poder ver el *prompt* completo que ha empleado el programa (Figura 13) es poder copiarlo y adaptarlo al gusto. La estructura ya se tiene, simplemente hay que corregir algunas palabras por otras para acercarse a lo que se busca.

Para la herramienta seleccionar, se despliega un pincel al cual se le puede variar el grosor. Con este, se pinta el elemento que se quiere editar y se escribe el comando en la caja de texto. Se realiza un ejemplo del proceso para transformar el rostro caucásico de la niña a uno asiático (Figura 14). La orden es: *asian girl face*.

Se aprecia en el renderizado final que es el instrumento idóneo para conservar exactamente la figura inicial.

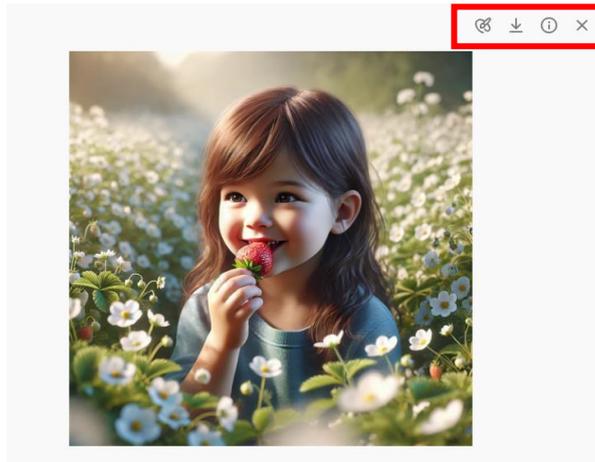


Figura 12. Funciones establecidas dentro de una imagen en DALL·E. Extraído de OpenAI. <https://openai.com/chatgpt/>

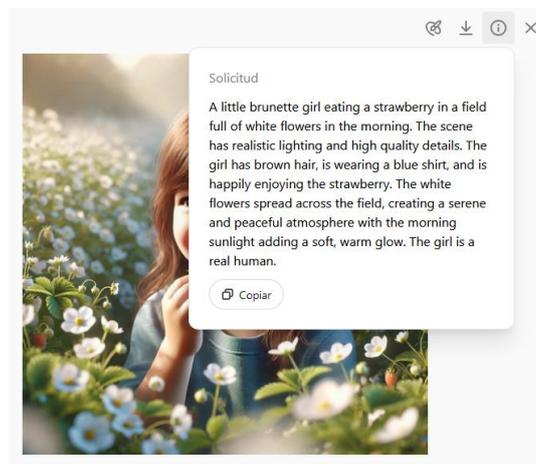


Figura 13. Visualización del *prompt* completo en una imagen en DALL·E. Extraído de OpenAI. <https://openai.com/chatgpt/>

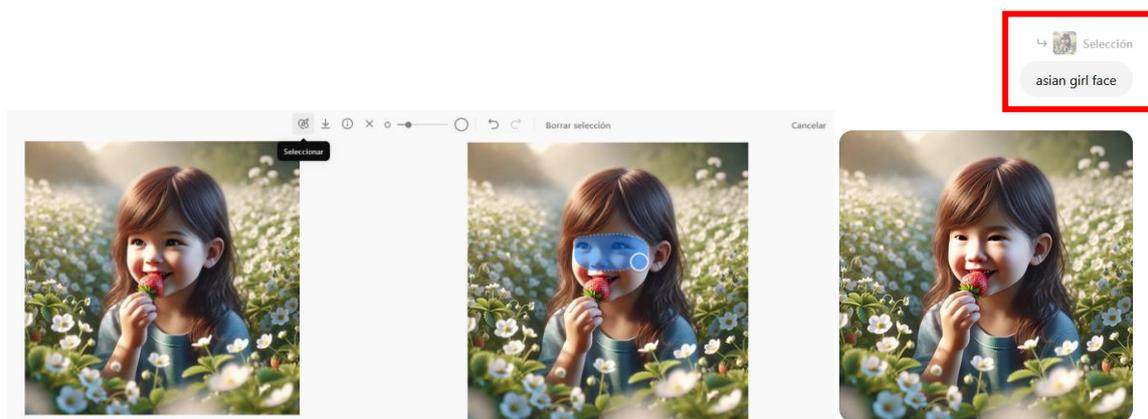


Figura 14. Proceso para transformar una imagen en DALL·E con la herramienta seleccionar. Elaboración propia a partir de OpenAI. <https://openai.com/chatgpt/>

Asimismo, DALL·E también permite cargar una fotografía personal desde el dispositivo que se esté utilizando y trabajar con ella.

3.2. Midjourney

- Acceso al programa:

Midjourney está implementado como bot¹⁰ desde la aplicación de Discord. De modo que, para poder manejarlo, se tiene que enlazar la cuenta de la plataforma con el plan de suscripción que se prefiera, puesto que, a diferencia del resto de modelos generativos, este no está abierto públicamente al uso gratuito.

El proceso de vinculación se efectúa automáticamente, pero previamente se deben seguir ciertos pasos para poder usarlo.

En primera instancia, hay que registrarse en Discord y cuando se acceda, en el menú lateral se desplegarán varias opciones, se clicará sobre la de crear un nuevo servidor y se confeccionará el nombre y otros parámetros (Figura 15).

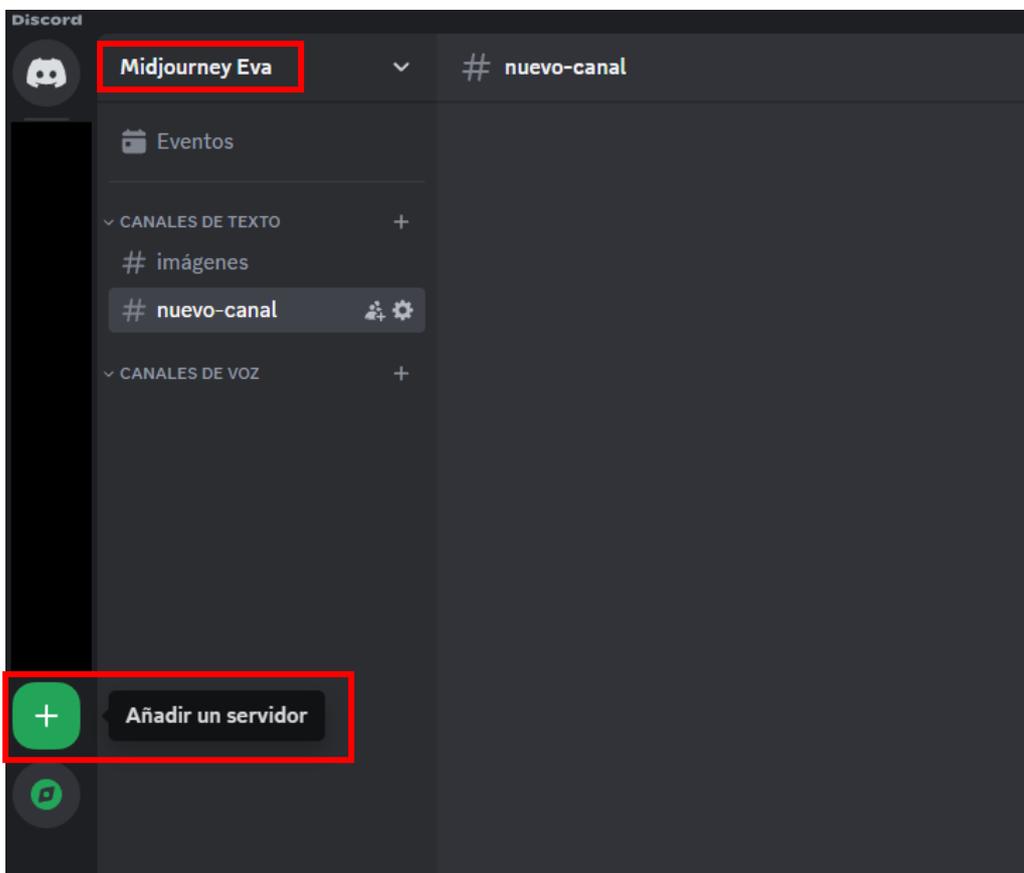


Figura 15. Creación de un servidor en Discord. Extraído de Discord. <https://discord.com/>

¹⁰ Aplicación informática que realiza tareas programadas desde Internet mediante unas instrucciones codificadas (AWS, 2024).

Luego, se prosigue con la opción de explorar servidores. Entonces, una barra de búsqueda aparecerá y se podrá teclear el de Midjourney (Figura 16).

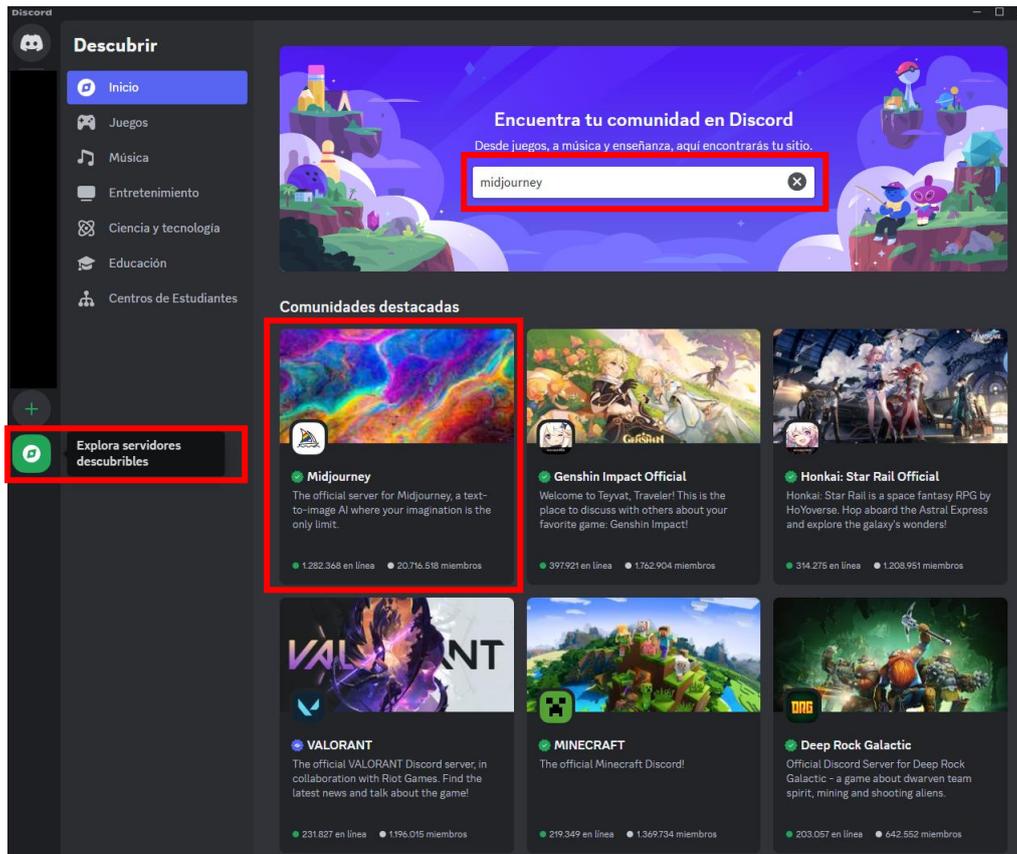


Figura 16. Menú de exploración de servidores de Discord. Extraído de Discord. <https://discord.com/>

Al ingresar en él, se contempla una comunidad que diseña y comparte en vivo todo tipo de imágenes, pero es importante agregar el bot que permite realizar esto en el servidor personal, ya que así se podrá disponer de forma privada y no se perderá el progreso que se lleve a cabo. Por ende: Canal newbie > Midjourney Bot > Perfil > Añadir aplicación > Añadir al servidor > Seleccionar servidor (Figuras 17 y 18).

Posteriormente, el programa enviará un mensaje con un enlace para comprar la suscripción que se precise en función de la utilidad que se le vaya a dar.

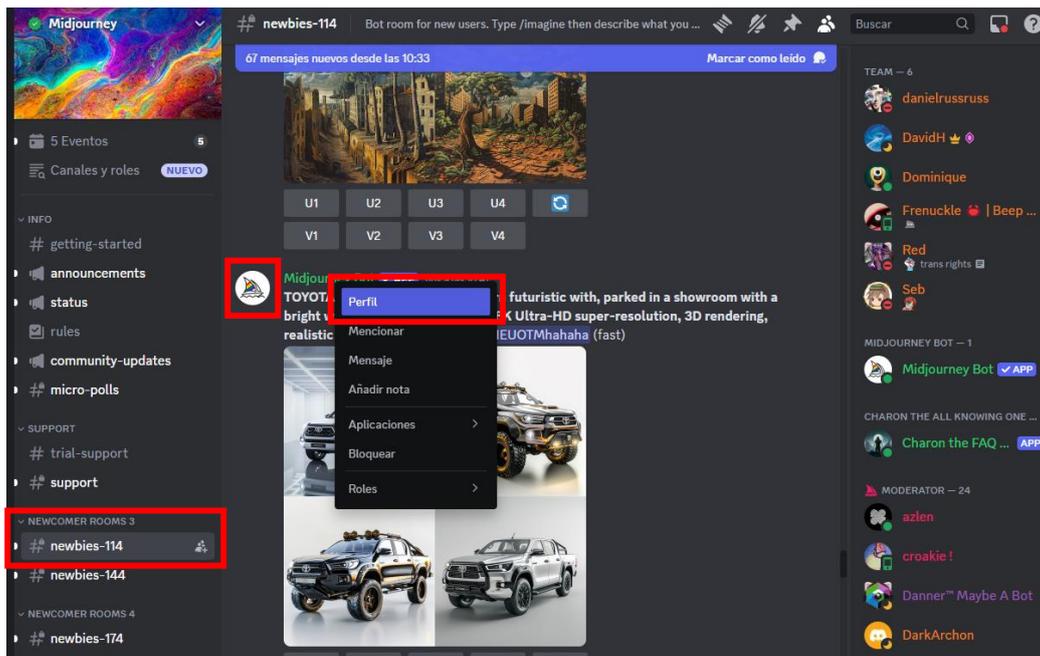


Figura 17. Proceso para implementar el bot Midjourney en un servidor personal de Discord (I). Extraído de Discord. <https://discord.com/>

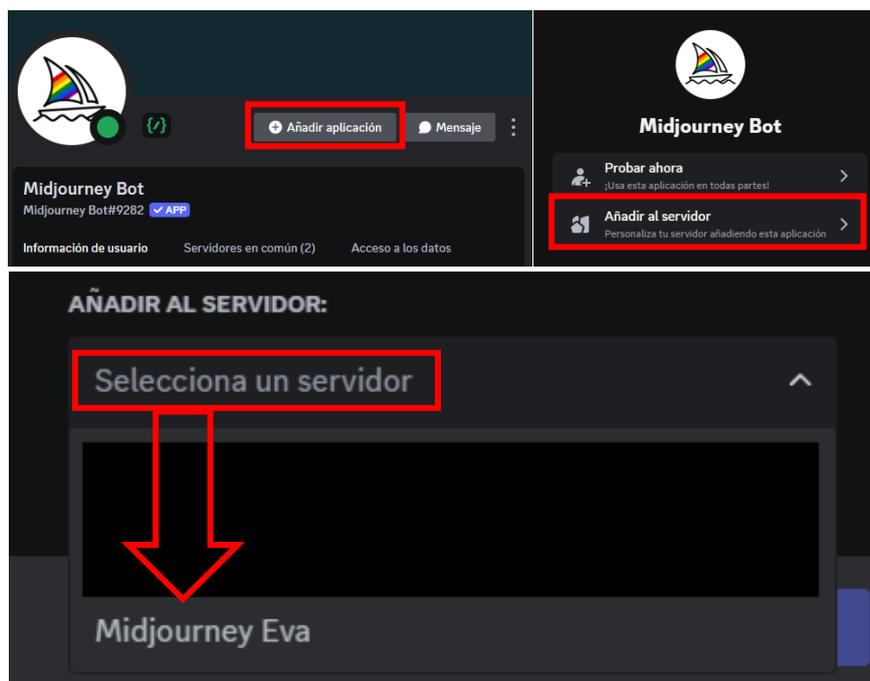


Figura 18. Proceso para implementar el bot Midjourney en un servidor personal de Discord (II). Extraído de Discord. <https://discord.com/>

Después de la configuración inaugural, ya se puede comenzar a producir conceptos.

- Generación de *prompt*:

Así pues, continuando con las mismas pautas que en DALL·E, acerca del idioma y la caracterización del *prompt*, se vuelve a emplear el equivalente aquí para observar la exactitud de los detalles.

Al contrario que los demás, en Midjourney no es suficiente con escribir el texto y procesar la información, sino que hay que utilizar un comando específico para cada tarea seguido de una barra (/).

Entre ellos se encuentran:

/settings: ajustar la configuración para personalizar el tiempo de renderizado, el nivel de estilización, el perfil público o privado y la resolución (activar siempre *Remix Mode* para cambiar los *prompt*)

/ask: recibir una respuesta a una pregunta

/blend: fusionar imágenes

/describe: escribir un *prompt* basado en un resultado

/help: ayuda al usuario

La función de generación viene constituida por */imagine*, de forma que la instrucción queda así: */imagine little brunette girl eating a strawberry in a field full of white flowers in the morning, realistic lighting, high quality* (Figura 19).

- Resultado:



Figura 19. Resultados de imagen estática generados por IA en Midjourney mediante un *prompt* de texto. Elaboración propia a partir de Discord. <https://discord.com/>

Las ilustraciones vienen determinadas por un aspecto de 1:1. En la parte inferior se distinguen diversas letras con números, haciendo referencia a la posición de la imagen, empezando de izquierda a derecha y de arriba abajo, y una rueda para crear cuatro resultados más. Las Us permiten editar el diseño con diferentes herramientas y descargar el archivo, mientras que las Vs facilitan la modificación el *prompt* de la figura elegida.

Por ejemplo, si se opta por la cuarta, se deberá seleccionar U4 o V4.

- Corrección:

Con la primera alternativa (U4) (Figura 20), se puede variar la solicitud de una manera más suave o más fuerte mediante la función *Vary*, dependiendo de la precisión que se quiera alcanzar. Para ello, basta con incorporar las palabras *wearing a blue shirt* en la línea preliminar (Figuras 21a y 21b).

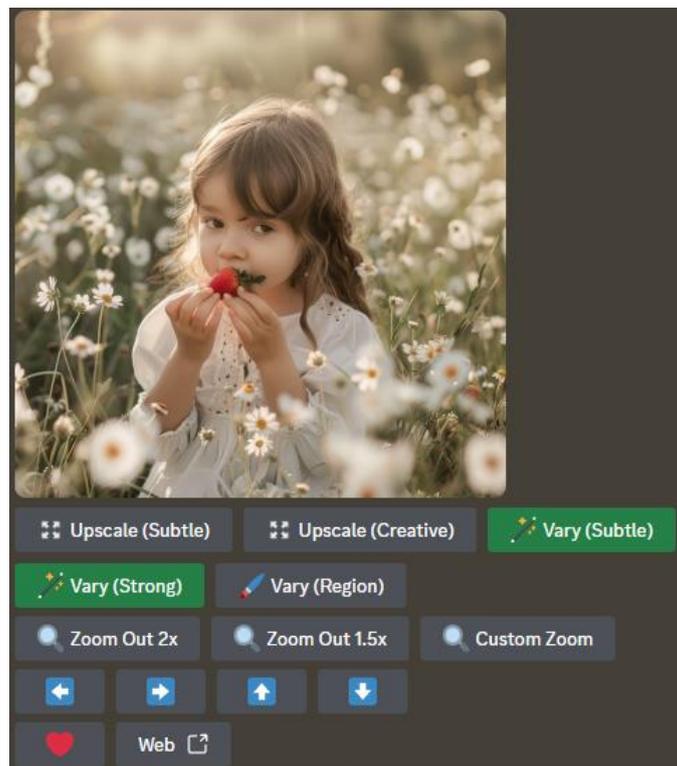


Figura 20. Resultados de imagen estática generados por IA en Midjourney modificando el *prompt* desde las funciones *Vary*. Elaboración propia a partir de Discord. <https://discord.com/>



Figuras 21a y 21b. Resultados de imagen estática generados por IA en Midjourney desde las funciones *Vary suave* (izq.) y *Vary fuerte* (dcha.). Elaboración propia a partir de Discord. <https://discord.com/>

Esta IA siempre ofrece cuatro respuestas independientemente de la transformación que se ejecute.

Conforme a las soluciones aportadas en ambos casos, se aprecia que la calidad y el estilo son idénticos. En cambio, ninguna de ellas conserva la niña original, aunque es verdad que la variación suave sí consigue mantener una postura semejante.

En cuanto a la segunda opción (V4), la de ajustar el comando directamente, se comprueba que el resultado es muy parecido al inicial porque sostiene unos parámetros similares (Figura 22), pero si el objetivo es generar el mismo sujeto con las correcciones aplicadas, tampoco es el instrumento indicado. En consecuencia, la única modalidad que cumple con esto es la de *Vary Region*, seleccionando la zona explícita e insertando la petición correspondiente (Figura 23).



Figura 22. Resultados de imagen estática generados por IA en Midjourney modificando el prompt original. Elaboración propia a partir de Discord. <https://discord.com/>

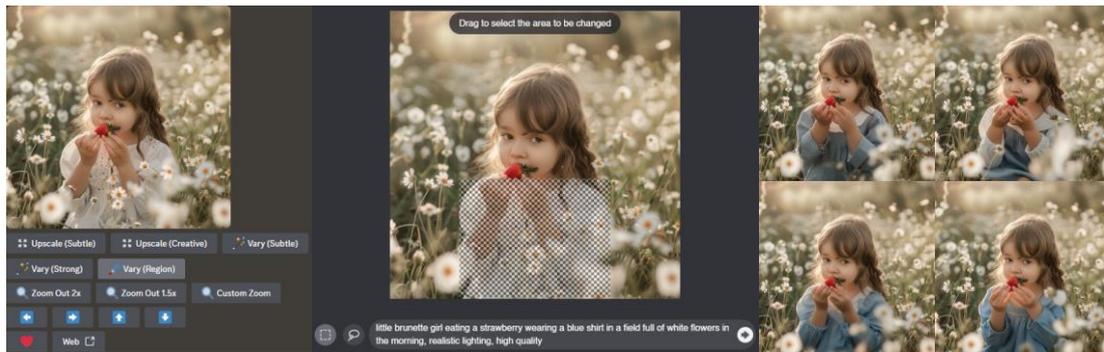


Figura 23. Proceso para transformar una imagen en Midjourney con la herramienta *Vary Region*. Elaboración propia a partir de Discord. <https://discord.com/>

Si se pretende adquirir otros efectos, como distintos estilos, ratios, enfoques, etc., será necesario incluir al final del *prompt* dos líneas (--) y el término oportuno para efectuar dicha función. Midjourney también es capaz de modelar figuras importadas desde un dispositivo.

Aquí se presentan algunas:

--ar: aspect ratio

--no: eliminación de componentes en la imagen

--tile: patrones para mosaicos

--seed: guarda los elementos de una imagen

En relación con los estilos, existen en la red infinidad de bibliotecas que recogen todo tipo de obras que sirven para inspirar a los usuarios, desde ambientaciones históricas, movimientos artísticos, ángulos de cámara, composiciones de color hasta técnicas de iluminación, entre otros (Figura 24).



Figura 24. Resultados de imagen estática generados por IA en Midjourney modificando el *prompt* original ahora con estilo basado en el s.XVII, pintura al óleo y con un ratio 3:2. Elaboración propia a partir de Discord. <https://discord.com/>

3.3. Leonardo AI

- Acceso al programa:

La compañía Leonardo AI brinda la oportunidad de acceder online a su propio modelo de generación desde la plataforma homónima. Se necesita un usuario y contraseña para iniciar sesión (Figura 25).

En la página principal, al igual que en el menú vertical, se hayan numerosas herramientas de las que se pueden hacer uso, no obstante, este caso se centra únicamente en la creación de imagen, en concreto, en el *text to image* (Figura 26).

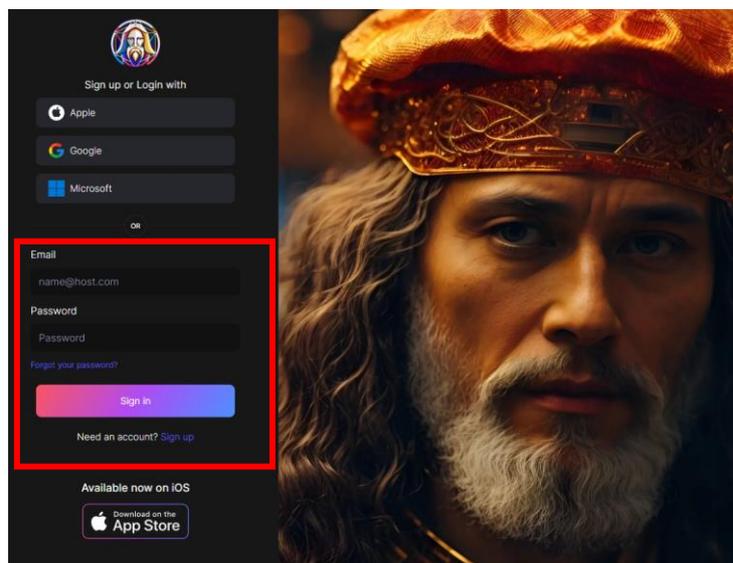


Figura 25. Página web principal de Leonardo AI donde se inicia sesión. Extraído de Leonardo AI. <https://leonardo.ai/>

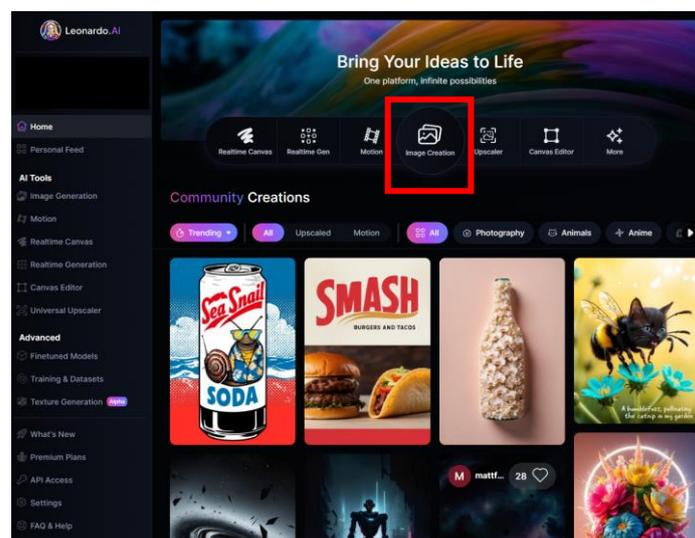


Figura 26. Menú principal de Leonardo AI con todas las herramientas disponibles. Extraído de Leonardo AI. <https://leonardo.ai/>

- Generación de *prompt*:

A causa de disponer exclusivamente de 150 tokens por la versión gratuita, el análisis se desempeñará con las funciones preestablecidas por el programa, en vista de que cada interacción consume entre 14 y 24 créditos y muchas otras ostentan un coste adicional.

La configuración es la siguiente:

- Contraste: medio
- Velocidad: rápida
- Dimensiones: 1:1
- Resolución: 1280x720
- Número de imágenes por procesamiento: 4

De todos los motores de renderizado que ofrece el sistema, se eligen los 3 que proporcionan un resultado más realista y potente, comenzando con *Leonardo Phoenix* y escribiendo la misma instrucción que van a tener en común los cuatro modelos estudiados en este TFG (Figura 27).

- Resultado:

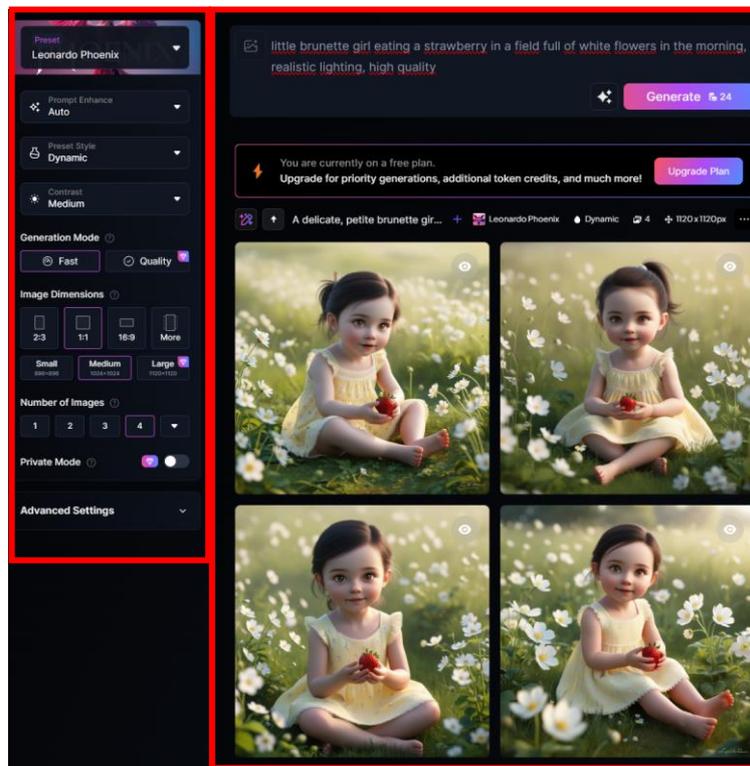


Figura 27. Configuración preestablecida y resultados de imagen estática generados por IA en Leonardo AI mediante un *prompt* de texto con el motor *Leonardo Phoenix*. Elaboración propia a partir de Leonardo AI. <https://leonardo.ai/>

Es el único generador que transforma el *prompt* dado en uno más extenso y complejo, el cual se puede visualizar desde la barra que se posiciona en la parte superior de las ilustraciones (Figura 28).

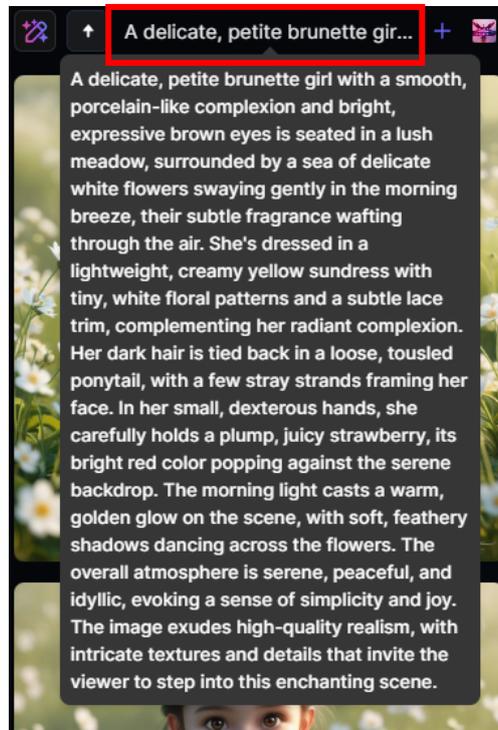


Figura 28. Visualización del *prompt* completo de una generación de imágenes en Leonardo AI. Extraído de Leonardo AI. <https://leonardo.ai/>

Por otra parte, los otros dos motores se describen como *Lifelike Vision* y *Stock Photography* (Figuras 29a y 29b), siendo este último el que se aproxima mejor a la realidad, de tal forma que sea con el que se concluya el resto de procedimiento.



Figuras 29a y 29b. Resultados de imagen estática generados por IA en Leonardo AI mediante un *prompt* de texto con el motor *Lifelike Vision* (izq.) y *Stock Photography* (dcha.). Elaboración propia a partir de Leonardo AI. <https://leonardo.ai/>

- Corrección:

Contrariamente a DALL·E y MidJourney, Leonardo AI no cuenta con una herramienta para modificar la solicitud desde el propio diseño ni tampoco con una selección que preserve el sujeto originario, solamente se puede formular de nuevo el comando. Se incluye una frase donde la niña lleva una camiseta azul (Figura 30).



Figura 30. Resultados de imagen estática generados por IA en Leonardo AI modificando el *prompt* original con el motor *Stock Photography*. Elaboración propia a partir de Leonardo AI. <https://leonardo.ai/>

Ahora, se le envía el mismo *prompt* que en Midjourney para transformar el estilo de pintura y ratio. Lamentablemente la respuesta no tiene efecto a excepción del dimensionado (Figura 31). Para ello, se debe escoger otro motor que lleve un estilo artístico programado, como el de anime, por ejemplo (Figura 32).



Figura 31. Resultados de imagen estática generados por IA en Leonardo AI modificando el *prompt* original con estilo basado en el s.XVII, pintura al óleo y con un ratio 16:9 (sin efecto en la respuesta). Elaboración propia a partir de Leonardo AI. <https://leonardo.ai/>



Figura 32. Resultados de imagen estática generados por IA en *Leonardo AI* modificando el prompt original con el motor *Anime*. Elaboración propia a partir de Leonardo AI. <https://leonardo.ai/>

En definitiva, esta IA muestra una elevada calidad en sus figuras, pero sufre de una gran desventaja al tener que recurrir a un motor en específico para crear un estilo artístico determinado, sin la opción de poder ejecutarlo desde la línea de comandos. Algunas funciones que paralelamente posee es la de bajar el fichero, borrar el fondo o cargar un retrato particular. Además, cuenta con un cuadro de texto para el *prompt* negativo (elementos a evitar en el diseño) y la posibilidad de entrenar un modelo propio.

3.4. Stable Difussion

- Acceso al programa:

Al ser Stable Difussion una IA con estructura de código abierto, se puede habilitar desde una aplicación local. Sin embargo, para facilitar el tener que descargarla y ajustarla dentro del ordenador, se utilizará la plataforma *online* Dream Studio. De esta manera, se compara así su alternativa en la red junto con los demás modelos que también operan conectados a Internet, salvo Midjourney.

No es obligatorio registrarse para acceder a la interfaz, pero sí para generar cualquier pieza visual (Figura 33).

- Generación de *prompt*:

Esta es la más simple de todas, en ella se identifica una pestaña para elegir el motor de renderizado, con la opción de descartarlo si no se prefiere trabajar con un estilo en especial, algo que en Leonardo AI no se puede llevar a cabo, un espacio para insertar tanto el *prompt* positivo como el negativo, una configuración para elegir la proporción de la imagen, el número de resultados por generación y la subida de un documento concreto (Figura 34).

Asimismo, cada archivo, puede guardarse y editarse con la función de selección, pintando el área en específico que se quiera cambiar y textualizando una nueva instrucción. Desafortunadamente, el diseño inicial no se conserva como se ha examinado en anteriores modelos.

Se empleará por defecto un ratio de 1:1 y 4 ilustraciones por procesamiento (con un valor de 0.91 créditos). Primero se usará el motor *Photographic* (Figura 35a), luego el *Cinematic* (Figura 35b) y por último se eliminarán (Figura 36). Dependiendo del que ofrezca la respuesta más realista, se establecerá para las siguientes modificaciones.

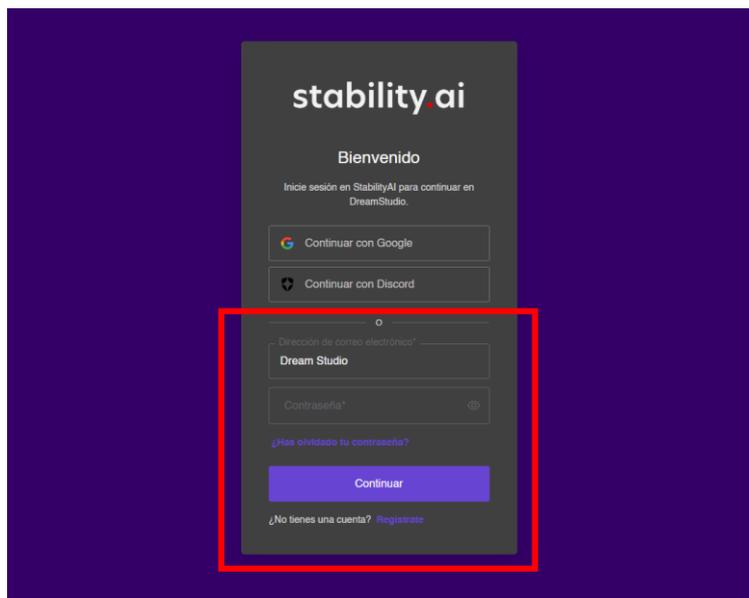


Figura 33. Página web principal de Dream Studio donde se inicia sesión. Extraído de Dream Studio. <https://beta.dreamstudio.ai/generate>

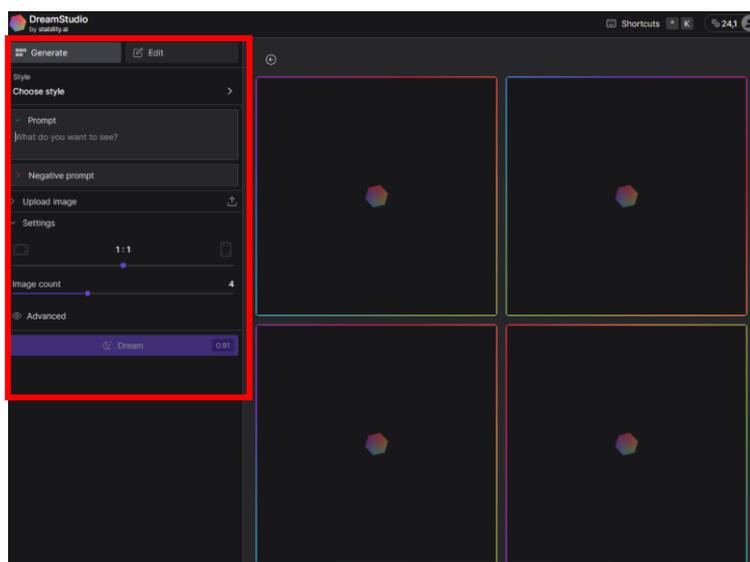


Figura 34. Interfaz de generación estática de Dream Studio. Extraído de Dream Studio. <https://beta.dreamstudio.ai/generate>

- Resultado:



Figuras 35a y 35b. Resultados de imagen estática generados por IA en Dream Studio mediante un *prompt* de texto con los motores de Stable Difussion *Photographic* (izq.) y *Cinematic* (dcha.). Elaboración propia a partir de Dream Studio. <https://beta.dreamstudio.ai/generate>



Figura 36. Resultados de imagen estática generados por IA en Dream Studio mediante un *prompt* de texto sin ningún motor de renderizado. Elaboración propia a partir de Dream Studio. <https://beta.dreamstudio.ai/generate>

Puesto que apenas se encuentra una diferencia muy clara, ya que la calidad que presentan es alta y los errores en los detalles son iguales, algunas manos están formadas por más de cinco dedos y diversos ojos se perciben borrosos, se escoge trabajar sin ningún motor para realizar las variaciones en la ropa (Figuras 37 y 38) y en la ambientación (Figura 39).

- Corrección:

No se admite corregir el comando desde la propia imagen, por lo que hay que volver a enviar la solicitud agregando los datos complementarios.



Figura 37. Resultados de imagen estática generados por IA en Dream Studio modificando el *prompt* original sin ningún motor de renderizado. Elaboración propia a partir de Dream Studio. <https://beta.dreamstudio.ai/generate>



Figura 38. Proceso para transformar una imagen en Dream Studio con la herramienta seleccionar. Elaboración propia a partir de Dream Studio. <https://beta.dreamstudio.ai/generate>



Figura 39. Resultados de imagen estática generados por IA en Dream Studio modificando el *prompt* original ahora con estilo basado en el s.XVII, pintura al óleo y con un ratio 16:9. Elaboración propia a partir de Dream Studio. <https://beta.dreamstudio.ai/generate>

3.5. Comparativa

En el análisis realizado, se ha demostrado el procedimiento introductorio que hay que realizar en cada uno de los cuatro sistemas optados para alcanzar sus relativas respuestas.

De modo que, para sintetizar visualmente los puntos más fuertes y débiles de cada uno, se ha construido una tabla con las características recogidas más importantes (Tabla 1).

Tabla 1. Comparación de los aspectos más relevantes entre DALL·E, Midjourney, Leonardo AI y Stable Difussion

Características relevantes	DALL·E	Midjourney	Leonardo AI	Stable Difussion
Coste de generación	Alto	Alto	Bajo	Bajo
Interfaz intuitiva	Alto	Bajo	Alto	Alto
Precisión de <i>prompt</i>	Alto	Alto	Alto	Alto
<i>Prompt</i> positivo	Alto	Alto	Alto	Alto
<i>Prompt</i> negativo	Bajo	Bajo	Alto	Alto
Imágenes por resultado	Bajo	Alto	Alto	Alto
Variaciones	Alto	Alto	Bajo	Bajo
Herramientas complementarias	Bajo	Alto	Alto	Bajo
Resultados realistas	Bajo	Alto	Alto	Alto
Calidad	Alto	Alto	Alto	Alto
Detalles sin errores	Bajo	Alto	Alto	Bajo
Tiempo de renderizado	Alto	Bajo	Bajo	Bajo
Comunidad activa y soporte	Bajo	Alto	Alto	Bajo

Leyenda	Bajo	Medio	Alto
----------------	-------------	--------------	-------------

Nota. Información extraída de la analítica previa de los cuatro modelos de IA generativa realizada en este TFG.

Por lo general, los modelos expuestos constituyen actualmente el panorama mundial de la IAG de imagen estática. Lo apreciado en la tabla es una aproximación a lo que se puede encontrar un usuario cuando navegue a través de la interfaz de alguno de ellos. Ninguno es perfecto, dado que la ventaja para uno supone la desventaja para otro. En cambio, generalmente el nivel de calidad es tan muy elevado.

En DALL·E y Stable Difussion, los detalles del físico de las personas (porque el de los objetos impecable), como las manos y los ojos podrían considerarse sus debilidades, ya que el ajuste de esos parámetros no es muy preciso y tampoco cuentan con apenas herramientas alternativas para ayudar en la labor de las conceptualizaciones, ni la comunidad es tan abierta para compartir soluciones.

Sin embargo, manifiestan uno de los tiempos de renderizado más rápido del sector, por encima de Midjourney y Leonardo AI.

Estos dos últimos poseen de los motores más eficaces y consistentes a la hora de crear una pieza creativa. Los elementos de la imagen cobran tal realismo que parecen sacados de un cuadro de museo o una fotografía profesional, los cuales muchos de ellos se pueden localizar en librerías y foros por todo Internet y redes sociales, gracias al gran volumen de internautas que apoyan este colectivo y dan soporte de forma global.

Existe un sinfín de estilos para probar y experimentar, por lo que cabe remarcar la fidelidad con la que elaboran las solicitudes basadas en toda clase de ambientaciones, épocas, colores, lugares, personajes, etc.

No obstante, Leonardo AI acarrea un enorme inconveniente al tener que ceñirse a sus motores estilizados para generar un diseño puntual, dejando al margen el texto que se indique en el *prompt*. El usuario ha de escoger entre los preestablecidos de la aplicación sin poder crear el suyo propio. Aunque, por otro lado, es el que más funciones contiene, algunas muy interesantes como la transformación *image to image* o los *Live Canvas*, que no se han podido profundizar debido a que se alejan de los límites de este proyecto.

En definitiva, cada uno de los cuatro presenta prestaciones variadas y no existe uno mejor. Dependerá del contexto en el que se vayan a utilizar, las necesidades que tenga la persona, la clase de propuesta que se quiera conseguir o el criterio adoptado.

Incluso en ocasiones, la combinación múltiple puede resultar muy útil, teniendo en cuenta que todos permiten cargar ficheros y trabajar con ellos. Lo que contribuye aún más a incrementar el desarrollo artístico en el sector audiovisual, y con la ayuda de las generaciones de imágenes tan fieles a sus solicitudes, garantizan poner en primer plano esta tecnología y su evolución.

4. CONCLUSIONES

En el presente Trabajo Final de Grado se fijaron al principio unos objetivos a cumplir que han servido de guía para desarrollar la siguiente memoria.

Como objetivo principal, se determinó la evaluación del funcionamiento y sistema de los modelos de Inteligencia Artificial Generativa de imagen estática de DALL·E, Midjourney, Leonardo AI y Stable Diffusion. Considero que la rigurosa investigación sobre el estudio puesto en práctica y la experiencia recabada al analizarlos en primer plano ha conferido su realización con éxito.

El proyecto ha sido desarrollado atendiendo a la tecnología situada hasta junio de 2024, de modo que las actualizaciones y la implementación de herramientas posteriores a esta fecha quedan excluidas del mismo.

Respecto a los objetivos secundarios, se han abordado en relación con el contexto histórico y social de la IA y el nacimiento de la generación de imágenes, desde los inicios hasta la actualidad, junto a su evolución en el sector audiovisual. Así como la arquitectura con la que están programadas y los diferentes tipos que existen, tarea vital para entender como están estructurados los modelos que se han explorado.

La repercusión que ha tenido el despegue del *text to image*, y sus posteriores consecuencias, dando pie a que cualquier persona con los conocimientos y equipo adecuados pueda crear un diseño en cuestión de segundos, también se ha definido en el recorrido que suscita la finalidad de esta reflexión.

Siguiendo la metodología establecida, se llega a cubrir la explicación de los aspectos fundamentales de las cuatro principales plataformas que integran las funciones de generación. Matiz que ha ayudado en el sucesivo análisis, en el cual se ha ahondado en cada una para conocer su interfaz, el modo de trabajo, la precisión con la que actúan los *prompt*, el descubrimiento de las comunidades que se hayan detrás, y, sobre todo, el nivel de calidad de las respuestas y el detalle realista de los resultados.

Por su parte, es conveniente mencionar algunas cuestiones éticas encontradas a lo largo del trabajo asociadas al uso de la IA generativa. Estas son la falta de transparencia en los algoritmos, la posible difusión de información falsa a partir de contenido generado con estas herramientas y la vulneración de los derechos de autor, entre otros. Todo ello enfatiza la urgencia de establecer normativas y marcos legislativos que regulen la práctica responsable de estos instrumentos y la necesidad de concienciar a la ciudadanía sobre un uso responsable y adecuado de estas herramientas.

Además, durante el análisis, se identificaron dificultades para proceder a su desempeño relacionadas con el acceso limitado a herramientas que requerían de licencias de pago o el coste de cada generación, lo que restringió la capacidad

de profundizar en ciertos aspectos técnicos y obtener otro tipo de resultados con los que formar una comparativa más amplia.

En cuanto a lo personal, creo que la elección de esta propuesta me ha despertado el interés en seguir profundizando en la materia, dado que me siento un poco más consolidada en la comprensión de esta tecnología. Espero también que este TFG pueda servir de manual inicial para acercar a estudiantes que quieran sumergirse en este fascinante mundo creativo.

5. BIBLIOGRAFÍA

- Alcaraz, I. (2023, 17 de julio). "La inteligencia artificial será la mayor burbuja de todos los tiempos", según el CEO de Stability AI. *El Confidencial*. Recuperado el 2 de julio de 2024, de https://www.elconfidencial.com/tecnologia/2023-07-17/inteligencia-artificial-burbuja-ceo-stability-ai_3701860/
- Alonso, R. (2024, 13 de marzo). Europa aprueba la primera ley de inteligencia artificial de la historia: estas son sus claves. *ABC Tecnología*. Recuperado el 15 de junio de 2024, de <https://www.abc.es/tecnologia/europa-aprueba-primer-ley-inteligencia-artificial-historia-20240313175642-nt.html?ref=https%3A%2F%2Fwww.abc.es%2Ftecnologia%2Feuropa-aprueba-primer-ley-inteligencia-artificial-historia-20240313175642-nt.html>
- Amazon Web Services. (2023). [Arquitectura interna de una GAN] [Imagen digital]. Amazon Web Services. Recuperado el 19 de junio de 2024, de <https://aws.amazon.com/es/what-is/gan/>
- Amazon Web Services. (2023). ¿Qué es un bot? *Amazon Web Services*. Recuperado el 8 de julio de 2024, de <https://aws.amazon.com/es/what-is/bot/>
- Amazon Web Services. (2023). ¿Qué es una GAN? *Amazon Web Services*. Recuperado el 19 de junio de 2024, de <https://aws.amazon.com/es/what-is/gan/>
- Barba, M. Á., & Valenzuela, I. (2023, 27 de julio). IA Generativa: Un cambio de paradigma para el que la sociedad y la industria deben estar preparados. *Telefónica*. Recuperado el 3 de junio de 2024, de <https://www.telefonica.com/es/sala-comunicacion/blog/ia-generativa-cambio-paradigma-sociedad-industria/>
- Bengesí, S., El-Sayed, H., Sarker, M. K., Houkpati, Y., Irungu, J., & Oladunni, T. (2024). Advancements in Generative AI: A comprehensive review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers. *IEEE Access*, 12(1), 69812-69837. <https://doi.org/10.1109/ACCESS.2024.3397775>
- Briggs, J., & Carnevali, L. (s. f.). Image-net: How ImageNet and AlexNet kickstarted the deep learning era of computer vision. *Embedding Methods for Image Search*. Pinecone. <https://www.pinecone.io/learn/series/image-search/imagenet/>

- Cabello, R. (2022, 22 de marzo). ¿Qué relación tienen los NFT con la IA? *Plain Concepts*. Recuperado el 22 de junio de 2024, de <https://www.plainconcepts.com/es/nft-inteligencia-artificial/>
- Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L., & Zimmel, R. (2023, 14 de junio). The economic potential of generative AI: The next productivity frontier. *McKinsey Digital*. Recuperado el 29 de junio de 2024, de <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-AIs-breakout-year>
- del Vayo, A. (2023, 29 de marzo). Leonardo AI es una inteligencia artificial alternativa a Midjourney con muchas más opciones. *El Español*. Recuperado el 1 de julio de 2024, de https://www.elespanol.com/elandroidelibre/noticias-y-novedades/20230329/leonardo-ai-inteligencia-artificial-alternativa-midjourney-opciones/752174827_0.html
- DeVon, C. (2023, 30 de noviembre). On ChatGPT's one-year anniversary, it has more than 1.7 billion users—here's what it may do next. *CNBC make it*. Recuperado el 3 de junio de 2024, de <https://www.cnbcm.com/2023/11/30/chatgpts-one-year-anniversary-how-the-viral-ai-chatbot-has-changed.html>
- Díaz, R. (2022, 16 de diciembre). Miles de artistas protestan en ArtStation contra las imágenes generadas por inteligencia artificial. *El Mundo*. Recuperado el 23 de junio de 2024, de <https://www.elmundo.es/tecnologia/creadores/2022/12/16/639c4306fdddff37f8b4595.html>
- Dircomfidencial. (2023, 12 de julio). Midjourney: ¿Qué es y cómo funciona? *Dircomfidencial*. Recuperado el 30 de junio de 2024, de <https://dircomfidencial.com/diccionario/midjourney-que-es-y-como-funciona-20230712-1337/>
- Discord. (2024). *Discord*. Recuperado el 8 de julio de 2024, de <https://discord.com/>
- DreamStudio (2024). *DreamStudio* (Versión del 10 de julio) [Large language model]. <https://beta.dreamstudio.ai/generate>
- Fernández, Y. (2023, 30 de enero). Stable Diffusion: qué es, qué puedes hacer con esta inteligencia artificial y maneras para usarlo en tu ordenador. *Xataka*. Recuperado el 2 de julio de 2024, de <https://www.xataka.com/basics/stable-diffusion-que-que-puedes-hacer-maneras-para-usarlo-tu-ordenador>
- Fernández, Y. (2023, 14 de septiembre). Qué es un prompt y por qué son tan importantes para usar la inteligencia artificial. *Xataka*. Recuperado el 19 de mayo de 2024, de <https://www.xataka.com/basics/que-prompt-que-importantes-para-usar-inteligencia-artificial>
- Fernández, Y. (2024, 19 de enero). Qué es MidJourney, cómo funciona y cómo puedes usar esta inteligencia artificial para crear imágenes. *Xataka*. Recuperado el 30 de junio de 2024, de <https://www.xataka.com/basics/que-midjourney>
- Fernández, Y. (2024, 16 de febrero). DALL-E 3: qué es, cómo funciona, sus novedades y qué puedes hacer con esta inteligencia artificial. *Xataka*. Recuperado el 29 de

- junio de 2024, de <https://www.xataka.com/basics/dall-e-3-que-como-funciona-sus-novedades-que-puedes-hacer-esta-inteligencia-artificial>
- Galán, J. S. (2020, 1 de agosto). Business to Business (B2B). *economipedia*. <https://economipedia.com/definiciones/business-to-business.html>
- García, L. (2023, 27 de septiembre). Fin a la huelga de guionistas de Hollywood: logros y causas pendientes. *elDiario.es*. Recuperado el 12 de junio de 2024, de https://www.eldiario.es/cultura/cine/huelga-guionistas-hollywood-logros-causas-pendientes_1_10548658.html
- Hidalgo, M. A. (2024, 22 de marzo). El papel de la percepción social en la implantación de la Inteligencia Artificial Generativa (I). *Nada Es Gratis*. <https://nadaesgratis.es/manu-hidalgo/el-papel-de-la-percepcion-social-en-la-implantacion-de-la-inteligencia-artificial-generativa-i>
- Himel, A. H. (2024, 25 de abril). Is Midjourney the best AI Image Generator? A Balanced Perspective [Entrada de blog]. *Vintagedomus*. Recuperado el 30 de junio de 2024, de <https://vintagedomus.com/is-midjourney-the-best-ai-image-generator-a-balanced-perspective/>
- Huang, J. (2024). API reference and guides (for developers). *Leonardo.Ai*. Recuperado el 1 de julio de 2024, de https://intercom.help/leonardo-ai/en/articles/8973587-api-reference-and-guides-for-developers#h_e182f07c12
- [Imagen digital de estructura de una red neuronal profunda en *Deep Learning*]. (s. f.). Xataka. Recuperado el 16 de junio de 2024, de <https://www.xataka.com/robotica-e-ia/machine-learning-y-deep-learning-como-entender-las-claves-del-presente-y-futuro-de-la-inteligencia-artificial>
- International Business Machines. (s. f.). What are large language models (LLMs)? *International Business Machines*. Recuperado el 29 de junio de 2024, de <https://www.ibm.com/topics/large-language-models>
- Lara, J. (2023). *Generación de imágenes con Inteligencia Artificial: revisión de la situación actual y aplicaciones prácticas* (Trabajo de Fin de Grado, Universidad Miguel Hernández). <https://hdl.handle.net/11000/30268>
- Lawlor, P., & Chang, J. (2024, 12 de febrero). The rise of generative AI: A timeline of breakthrough innovations. *Qualcomm*. Recuperado el 3 de junio de 2024, de <https://www.qualcomm.com/news/onq/2024/02/the-rise-of-generative-ai-timeline-of-breakthrough-innovations>
- Leonardo.Ai. (2024). *Leonardo.Ai*. Recuperado el 1 de julio de 2024, de <https://leonardo.ai/>
- Leonardo.Ai. (2024). *Leonardo.Ai* (Versión del 10 de julio) [Large language model]. <https://app.leonardo.ai/auth/login?callbackUrl=%2F>
- Leonardo.ai. (2024). En PitchBook [Base de datos]. Recuperado el 1 de julio de 2024, de <https://pitchbook.com/profiles/company/530626-78#overview>
- Lin, L. (2023). Standing on the Shoulders of AI Giants. *IEEE Computer Society*, 56(1), 97-101. <https://doi.ieeecomputersociety.org/10.1109/MC.2022.3218176>

- McCarthy, J. (2007). *What is artificial intelligence?* Stanford University. Recuperado el 29 de mayo de 2024, de <https://www-formal.stanford.edu/jmc/whatisai.pdf>
- Merritt, R. (2022, 19 de abril). ¿Qué es un modelo transformer? [Entrada de blog]. Blog Nvidia. Recuperado el 20 de junio de 2024, de <https://la.blogs.nvidia.com/blog/que-es-un-modelo-transformer/>
- Midjourney. (2024). *Midjourney*. Recuperado el 30 de junio de 2024, de <https://www.midjourney.com/home>
- Midjourney. (2024). *Midjourney*. (Versión del 8 de julio) [Large language model]. <https://www.midjourney.com/home>
- MIT Technology. (2023, 18 de julio). The great acceleration: CIO perspectives on generative AI. *MIT Technology Review*. Recuperado el 21 de junio de 2024, de <https://www.technologyreview.com/2023/07/18/1076423/the-great-acceleration-cio-perspectives-on-generative-ai/>
- National Cybersecurity Alliance (2023, 22 de diciembre). Cómo protegerse contra los deepfakes. *National Cybersecurity Alliance*. Recuperado el 23 de junio de 2024, de <https://staysafeonline.org/es/resources/how-to-protect-yourself-against-deepfakes/>
- OpenAI. (2024). *DALL·E*. (Versión del 5 de julio) [Large language model]. <https://chat.openai.com>
- OpenAI. (2024). *DALL·E 3*. Recuperado el 29 de junio de 2024, de <https://openai.com/index/dall-e-3/>
- Página12. (2023, 29 de marzo). Artistas en lucha contra los programas de IA que copian sus estilos. *Página12*. Recuperado el 12 de junio de 2024, de [https://www.pagina12.com.ar/535546-artistas-en-lucha-contra-los-programas-de-ia-que-copian-sus-](https://www.pagina12.com.ar/535546-artistas-en-lucha-contra-los-programas-de-ia-que-copian-sus-de-ia-que-copian-sus-)
- Palomino, E. (2024, 2 de abril). Billie Eilish, Nicki Minaj y más artistas firmaron carta contra la IA que “sabotea la creatividad”. *Infobae*. Recuperado el 12 de junio de 2024, de <https://www.infobae.com/entretenimiento/2024/04/02/billie-eilish-nicki-minaj-y-mas-artistas-firmaron-carta-contra-la-ia-que-sabotea-la-creatividad/>
- Pereyra, W. (2022, 10 de mayo). ¿Qué es Blockchain o Tecnología de Registro Distribuido (DLT)? [Entrada de blog]. Blog Posgrado - Universidad Continental. Recuperado el 22 de junio de 2024, de [https://blogposgrado.ucontinental.edu.pe/blockchain-tecnologia-registro-distribuido-dlt#:~:text=Distribuido%20\(DLT\).-Una%20tecnolog%C3%ADa%20de%20registro%20distribuido%20almacena%20la%20informaci%C3%B3n%20en%20m%C3%BAltiples,para%20que%20la%20descentralizaci%C3%B3n%20funcione](https://blogposgrado.ucontinental.edu.pe/blockchain-tecnologia-registro-distribuido-dlt#:~:text=Distribuido%20(DLT).-Una%20tecnolog%C3%ADa%20de%20registro%20distribuido%20almacena%20la%20informaci%C3%B3n%20en%20m%C3%BAltiples,para%20que%20la%20descentralizaci%C3%B3n%20funcione)
- Polo, D., Micaletto, J. P., & Martín, P. (2023). Impacto y oportunidades de la inteligencia artificial en el arte visual: personalización de la creatividad con Stable Diffusion y Controlnet. En M. C. Gálvez, M. C. Casado & B. Ventura-Salom (Coords.), *Comunicación, creación artística y audiovisual: un marco para la innovación educativa*. (pp. 894-909). Dykinson.

<https://www.researchgate.net/publication/375422553> Impacto y oportunidades de la inteligencia artificial en el arte visual personalización de la creatividad con Stable Diffusion y Controlnet

Raza, A. (2024, 2 de febrero). Unveiling the complete history of Leonardo AI: From concept to creation. *Medium*. Recuperado el 1 de julio de 2024, de <https://medium.com/@aliraza/unveiling-the-complete-history-of-leonardo-ai-from-concept-to-creation-f1ec17cb48de>

Recreación de Obama y Trump jugando al baloncesto [Imagen digital]. (2023). MDTech. Recuperado el 23 de junio de 2024, de <https://www.mdtech.news/servicios/2023/3/24/esta-ia-recreo-donald-trump-barack-obama-jugando-al-baloncesto-8456.html>

Rico, J. (2023). *Nuevos retos para el diseño y la comunicación. La inteligencia artificial en los procesos creativos del diseño gráfico*. (Tesis doctoral, Universitat Politècnica de València). <https://doi.org/10.4995/thesis/10251/192876>

Rodríguez, T. (2020, 16 de octubre). Machine Learning y Deep Learning: cómo entender las claves del presente y futuro de la inteligencia. *Xataka*. Recuperado el 16 de junio de 2024, de <https://www.xataka.com/robotica-e-ia/machine-learning-y-deep-learning-como-entender-las-claves-del-presente-y-futuro-de-la-inteligencia-artificial>

Rosso, C. (2019, 10 de enero). What caused the AI renaissance: Much ado about Deep Learning Backpropagation. *Psychology Today*. Recuperado el 30 de mayo de 2024, de <https://www.psychologytoday.com/us/blog/the-future-brain/201901/what-caused-the-ai-renaissance>

Sainza, B. (2023, 25 de octubre). AI, NFT: El papel del artista con la llegada de las nuevas tecnologías. *Experimenta*. Recuperado el 22 de junio de 2024, de <https://www.experimenta.es/noticias/grafica-y-comunicacion/ai-nft-el-papel-del-artista-con-la-llegada-de-las-nuevas-tecnologias/>

Shu, C. (2023, 6 diciembre). Sydney-based generative AI art platform Leonardo.AI raises \$31M. *TechCrunch*. Recuperado el 1 de julio de 2024, de https://techcrunch.com/2023/12/06/leonardo-ai/?gucounter=1&gucerefefferer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&gucerefefferer_sig=AQAAAGbDRReytRXSqiycdEOnOpqWwJNILw8ScGS2zmEMnXWRL6kILVPVrpL85jP8W322DhYGpzA6sOhH4_dFHgGDEPbyMcaeCzRyAIJ_e-BvsQdU7uKIXJnvLkfF2o5aTDdJ07J0liJjacHwO1I0hyttK4ffMVf-dpXk4NmnYUealJ0Qw

Singla, A., Sukharevsky, A., Hall, B., Yee, L., & Chui, M. (2023, 1 de agosto). The state of AI in 2023: Generative AI's breakout year. *QuantumBlack AI by McKinsey*. Recuperado el 10 de junio de 2024, de <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-AIs-breakout-year>

Stability.AI. (2024). *Stability.AI*. Recuperado el 2 de julio de 2024, de <https://stability.ai/about>

- Ulgus-Servant, V. (2023, 9 de noviembre). Fin de la huelga en Hollywood: un acuerdo histórico entre el sindicato de actores y los estudios. *Vanity Fair*. Recuperado el 12 de junio de 2024, de <https://www.revistavanityfair.es/articulos/fin-huelga-actores-guionistas-hollywood-acuerdo>
- Wiggers, K., & Coldewey, D. (2024, 16 de marzo). This week in AI: Midjourney bets it can beat the copyright police. *TechCrunch*. Recuperado el 30 de junio de 2024, de <https://techcrunch.com/2024/03/16/this-week-in-ai-midjourney-bets-it-can-beat-the-copyright-police/>
- Xavier, P. (2023). *El deep fake del Papa Francisco con una chaqueta Balenciaga* [Imagen digital]. La Vanguardia. Recuperado el 23 de junio de 2024, de <https://www.lavanguardia.com/tecnologia/20230328/8859266/creador-imagen-viral-papa-francisco-chaqueta-deep-fake-pmv.html>
- Yi, Q., Chen, X., Zhang, C., Zhou, Z., Zhu, L., & Kong, X. (2024). Diffusion models in text generation: a survey. *PeerJ Computer Science*, 1(1), 1-26. <https://doi.org/10.7717/peerj-cs.1905>
- Zhang, L., Rao, A., & Agrawala, M. (2023). Adding conditional control to Text-to-Image Diffusion Models. *IEEE/CVF International Conference on Computer Vision (ICCV)*, 1(1), 3813-3824. <https://doi.org/10.1109/ICCV51070.2023.00355>