DALL-E

Anton Wallentin Delgado

Uno de los grandes avances en la tecnología de la IA viene siendo, DALL-E, el cual es un sistema de inteligencia artificial desarrollado por OpenAI, que a diferencia de ChatGPT, el cual se basa en el procesamiento del lenguaje natural (NLP) y es un modelo de lenguaje generativo, lo que significa que esta diseñado para entender y generar texto en respuesta a las entradas de texto que recibe, DALL-E se basa en modelos de lenguaje grandes (LLM) multimodal. DALL-E puede entender, interpretar y generar o actuar sobre múltiples tipos de datos de entrada, como texto, imágenes, audio y video [1]. Este sistema toma su nombre de la combinación del famoso pintor Salvador Dalí y el personaje WALL-E de Pixar, reflejando sus capacidades artísticas y su naturaleza de IA. DALL-E y tiene como propósito principal cerrar la brecha entre la comprensión del lenguaje natural y la representación visual, permitiendo a los usuarios describir imágenes deseadas en texto y generando visuales que coincidan con esas descripciones. Lo que lo hace notable es su capacidad para converger los campos del modelado de lenguaje e imagen en una manera sin precedentes, brindando una visión del futuro del contenido visual generado por IA. DALL-E utiliza un modelo de aprendizaje profundo basado en GPT-3, destacado por su capacidad de comprensión del lenguaje natural, y emplea una variante de la arquitectura “Transformer”. Se entrenó con un vasto conjunto de datos de pares de texto e imagen de internet, aprendiendo a asociar descripciones textuales con representaciones visuales correspondientes. A diferencia de los modelos de generación de imágenes tradicionales, DALL-E puede producir una amplia gama de imágenes basadas en el texto proporcionado, mostrando un nivel impresionante de generalización y creatividad. Las aplicaciones de DALL-E son variadas, incluyendo diseño gráfico, publicidad, juegos, entretenimiento, comercio electrónico, educación, investigación y arte. Sin embargo, todo no es perfecto, DALL-E enfrenta desafíos como la generación coherente de imágenes, control de calidad de imagen, sesgos en los conjuntos de datos, cuestiones de derechos de autor y altos requerimientos computacionales. Además, existen limitaciones en su capacidad para generar imágenes altamente detalladas y la consistencia en la generación de imágenes basada en variaciones textuales mínimas. También se plantean preocupaciones éticas relacionadas con la posibilidad de generar contenido inapropiado o dañino y el impacto en los trabajos en la industria creativa [2]. Indagando más sobre este último punto, las opiniones están muy divididas ya que algunos ven esta tecnología como una herramienta útil que puede fomentar la creatividad, mientras que otros temen por la pérdida de sus empleos. Además, se discute la cuestión de la propiedad intelectual de las imágenes generadas por IA, planteando preguntas sobre quién posee los derechos de autor de tales imágenes: si los programadores detrás de la tecnología, los usuarios que proporcionan las indicaciones, o una combinación de ambos [3]. Aunque enfrenta obstáculos y restricciones, el porvenir de DALL-E se muestra alentador, abriendo el camino a avances en la creación de imágenes y la cooperación entre diversas disciplinas para ampliar su uso en el mundo real. DALL-E se destaca como un modelo ejemplar y vanguardista en la tecnología de conversión de texto a imagen mediante inteligencia artificial, ofreciendo amplias oportunidades para transformar la manera en que producimos y adaptamos contenido visual.

Para culminar, añadí una última referencia sobre unos datos estadísticos sobre la IA que me parecieron muy interesantes y de suma importancia [4].

# Referencias

|  |  |
| --- | --- |
| [1] | T. Filzinger, «Konfuzio,» 19 October 2023. [En línea]. Available: https://konfuzio.com/en/multimodal-llm/. [Último acceso: 2 2 2024]. |
| [2] | R. Schmidt, «AppMaster,» 6 November 2023. [En línea]. Available: https://appmaster.io/blog/challenges-and-limitations-dall-e. [Último acceso: 2 2 2024]. |
| [3] | D. P. &. S. R. STROUD, «Media Ethics,» 2022. [En línea]. Available: https://www.mediaethicsmagazine.com/index.php/browse-back-issues/219-fall-2022-vol-34-no-1/3999403-an-extension-of-the-artist-dall-e-2-and-the-ethical-challenges-of-ai-art. [Último acceso: 2 2 2024]. |
| [4] | K. H. &. R. Watts, «Forbes,» 25 April 2023. [En línea]. Available: https://www.forbes.com/advisor/business/ai-statistics/#top\_ai\_statistics\_section. [Último acceso: 2 2 2024]. |