



Universidad Pontificia de Comillas, ICADE

GENERACIÓN DE PICTOGRAMAS CON INTELIGENCIA ARTIFICIAL GENERATIVA PARA MIEMBROS DEL ESPECTRO AUTISTA

Autor: Clara Maldonado Gilarranz

Director: Eduardo César Garrido Merchán

Clave: 201910755

MADRID | Abril 2024

RESUMEN

El presente proyecto de investigación se centra en el uso de la Inteligencia Artificial Generativa (IAG) para mejorar la comunicación de personas con dificultades verbales, especialmente para miembros del espectro autista. Se presenta una arquitectura de red neuronal convolucional diseñada específicamente para la clasificación de imágenes, con el objetivo de evaluar la efectividad de los pictogramas generados por Inteligencia Artificial (IA) en comparación con los pictogramas convencionales. Tras una rigurosa experimentación se observa que, si bien los pictogramas generados por IA muestran una calidad y utilidad comparables a los de fuentes convencionales, como es ARASAAC, aún presentan diferencias significativas. Aunque representan una alternativa innovadora, se hace evidente la necesidad de un refinamiento para lograr una mayor similitud con los pictogramas originales y, por consiguiente, una efectividad superior en la comunicación asistida. La tesis que se presenta establece un marco metodológico para la evaluación de pictogramas generados por IA, lo que puede servir como punto de partida para futuras investigaciones en el campo de la generación de imágenes, con el propósito de mejorar la calidad y la accesibilidad de las herramientas de comunicación para personas con dificultades del habla.

Palabras clave: Espectro autista, Inteligencia Artificial Generativa (IAG), Inteligencia Artificial (IA), Comunicación, Pictogramas, Modelo, Dificultades del habla, Imágenes, Red neuronal convolucional.

ABSTRACT

The present research project focuses on the use of Generative Artificial Intelligence (GAI) to enhance communication for individuals with verbal difficulties, particularly for those on the autism spectrum. It introduces a convolutional neural network architecture specifically designed for image classification, aiming to assess the effectiveness of Artificial Intelligence (AI) generated pictograms compared to conventional ones. Through rigorous experimentation, it is observed that while AI-generated pictograms demonstrate quality and utility comparable to those from conventional sources like ARASAAC, they still exhibit significant differences. Despite representing an innovative alternative, the need for refinement becomes evident to achieve greater resemblance to original pictograms and, consequently, higher effectiveness in assisted communication. The thesis at hand establishes a methodological framework for evaluating AI-generated pictograms, which could serve as a starting point for future research in the field of image generation, aiming to improve the quality and accessibility of communication tools for individuals with speech difficulties.

Keywords: *Autistic spectrum, Generative Artificial Intelligence (GAI), Artificial Intelligence (AI), Communication, Pictograms, Model, Speech difficulties, Images, Convolutional neural network.*

Índice de contenido

1. INTRODUCCIÓN.....	6
2. ESTADO DEL ARTE.....	10
3. ALCANCE DE LA TESIS DE FIN DE GRADO	16
4. DESCRIPCIÓN DEL MODELO.....	18
5. DESCRIPCIÓN DEL PROBLEMA.....	22
6. EXPERIMENTOS Y RESULTADOS	26
7. COMERCIALIZACIÓN DE LA SOLUCIÓN.....	28
8. CONCLUSIONES.....	32
9. FUTURAS LÍNEAS DE INVESTIGACIÓN	34
10. REFERENCIAS BIBLIOGRÁFICAS	36

Índice de figuras

Ilustración 1. Convolución, Calvo, 2018.....	20
Ilustración 2. Red Convolutiva, Costa, 2020.....	21
Ilustración 3. Diagrama de flujo. Elaboración propia.	31

Índice de tablas

Tabla 1. Estructura del Estado del Arte. Elaboración propia.	10
---	----

Índice de ecuaciones

Ecuación (1): Operación de convolución	19
Ecuación (2): Tamaño de salida de capa de agrupación	20

1. INTRODUCCIÓN

La contextualización del nacimiento y desarrollo de la Inteligencia Artificial (IA) se comprende mejor observando sus orígenes históricos. Hay un acuerdo general sobre el hecho de que la IA surgió en la década de los cincuenta, con la investigación de programas informáticos simbólicos para la resolución automatizada de problemas (Amador Hidalgo, 1996). Sin embargo, es esencial retroceder aún más en el tiempo para encontrar la aparición de las primeras ideas y conceptos sobre la IA, los cuales se remontan al siglo XIX. Es destacable mencionar que figuras como Ada Lovelace fueron precursoras en este ámbito, al programar el primer algoritmo para máquinas, lo que anticipó la idea revolucionaria de que las máquinas podrían llevar a cabo cálculos generales (Abeliuk y Gutiérrez, 2021). Se observa también que, desde sus inicios, los progresos en el área de los sistemas inteligentes han estado estrechamente vinculados con los desarrollos alcanzados en el campo de la informática (Amador Hidalgo, 1996). Lovelace estableció así los fundamentos para lo que posteriormente se convertiría en la disciplina de la IA, resaltando la noción de que las máquinas podrían desempeñar funciones que trascendieran las meras operaciones numéricas (Abeliuk y Gutiérrez, 2021).

Unos años más tarde, en el siglo XX, surgen figuras como el matemático Alan Turing, una figura que contribuyó notablemente al surgimiento y a la consolidación de la IA (Lombana Bermudez, 2018). Entre sus numerosas contribuciones al campo se encuentra su influyente obra titulada “Computing Machinery and Intelligence” (Turing, 1950). En este trabajo seminal, Turing planteó la interrogante fundamental sobre la capacidad de las máquinas para pensar, anticipando así debates y controversias que eventualmente darían lugar al famoso Test de Turing (Lombana Bermudez, 2018). Dicho test, diseñado para evaluar la IA de una máquina en comparación con la humana, ha sido crucial en el desarrollo y la evaluación de futuros sistemas de modelos largos del lenguaje (Lombana, 2018). Pocos años más tarde se formalizaría la IA como nuevo campo de estudio científico en la conferencia de Darmouth (Abeliuk y Gutiérrez, 2021). Estos hitos históricos marcan el comienzo de una disciplina que ha pasado por múltiples fases de desarrollo a lo largo de la historia. Con el transcurso del tiempo, la IA ha experimentado una evolución continua, extendiéndose a nivel global y siendo implementada en variados sectores, tales como las finanzas y la educación. Su notable

adaptación y expansión subraya la creciente presencia actual de la IA en aspectos fundamentales de la sociedad y la economía.

Al explorar esta trayectoria histórica de la IA, se hace evidente una transformación notable en las formas de interacción, tanto entre individuos como entre personas y máquinas. Dentro del vasto ámbito de la comunicación, nos encontramos múltiples formas de interacción que trascienden o van más allá de la expresión verbal. Además, existen ciertos colectivos que presentan dificultades o que directamente no pueden recurrir a la comunicación verbal. Para estos determinados grupos, como individuos con trastornos del espectro autista o afectados por discapacidades del habla, la comunicación no verbal adquiere una relevancia significativa. En este contexto, la incapacidad para emplear el lenguaje oral restringe su capacidad para expresarse y comprender el entorno que les rodea, lo que hace que la comunicación no verbal surja como un recurso esencial y muy valioso.

Por ende, el problema que este trabajo se propone abordar es una de las modalidades de comunicación disponibles para aquellos colectivos que enfrentan dificultades en el ámbito verbal. Dicha vía de interacción son los pictogramas, representaciones visuales de imágenes o símbolos diseñados para ser fáciles de comprender y simplificar las interacciones, sin depender del lenguaje escrito. El problema reside en que los pictogramas que se recogen para la comunicación de los distintos grupos son tan limitados como fijos. Dicha limitación en la diversidad de los pictogramas supone un obstáculo para la expresión completa y precisa de los colectivos. De esta forma, se vuelve necesario recurrir al paradigma de la Inteligencia Artificial Generativa (IAG), un subcampo de la IA que se enfoca en la creación de contenido a partir de datos preexistentes como texto. Además, la IAG ya ha demostrado su habilidad en la creación de imágenes complejas, lo que la convierte en una potencial solución para abordar esta dificultad (Franganillo, 2023).

El objetivo principal de este proyecto es crear pictogramas que sean más versátiles, instantáneos y que no requieran realizar una búsqueda visual para encontrarlos, todo ello a través del uso de la IAG con *prompts*. La solución propuesta implica la utilización de la avanzada herramienta DALLE-3 a través de la plataforma de ChatGPT. Empleando esta herramienta se podrá crear un chat personalizado para la generación de

imágenes basadas en descripciones textuales. La elección de esta herramienta se debe a su capacidad para adaptarse perfectamente a la generación de representaciones visuales dinámicas e inmediatas, lo que la hace idónea para cumplir con los requisitos del proyecto.

El presente documento se estructura en siete partes principales. Comienza con la exposición del estado de la cuestión, donde se lleva a cabo una exhaustiva revisión de la literatura existente sobre los modelos de *text-to-image* y el uso de la IAG en relación al colectivo del espectro autista. A continuación, se encuentra el alcance de la Tesis de Fin de Grado (TFG) que comprende un detallado listado de objetivos, hipótesis, asunciones y restricciones del proyecto. Posteriormente, se detalla la metodología utilizada para llevar a cabo el análisis, describiendo el modelo empleado para desarrollar el trabajo. El desarrollo de la investigación se divide en varias secciones fundamentales que incluyen la justificación del modelo elegido, el proceso de entrenamiento del modelo, su posterior refinamiento y la implementación de un clasificador de imágenes. Seguidamente, se presenta un resumen detallado de los experimentos realizados y los resultados obtenidos a lo largo del estudio, seguido de las conclusiones relativas al estudio. Finalmente, se incluyen las referencias bibliográficas que respaldan el estudio en orden alfabético, además de incluir también el último acceso para aquellas fuentes que contengan un enlace en línea, asegurando así que se proporcione el contexto temporal de la información consultada.

2. ESTADO DEL ARTE

El siguiente estado del arte se organiza en tres partes principales que abordan diferentes aspectos de la intersección entre la IA y la comunicación en el contexto de la discapacidad. En primer lugar, se explora el impacto de los modelos fundacionales en general en la creación de contenido, con un enfoque especial en la generación de imágenes a partir de descripciones textuales. En segundo lugar, se aborda la discapacidad del habla, particularmente en el contexto del trastorno del lenguaje y del espectro autista. Y, por último, se investiga el papel de la IA en la mejora de la comunicación de los individuos con discapacidad no verbal.

Secciones:	Referencias principales
Generación de contenido a través de IA, especialmente la creación de imágenes.	<ul style="list-style-type: none">• Franganillo, 2023.• Peñalvo et al., 2024.• Zhang et al., 2023.
Análisis de la discapacidad del habla, en el contexto del trastorno del lenguaje y del espectro autista.	<ul style="list-style-type: none">• Artigas-Pallarès et al., 2022.• Rupérez et al., 2022.
IA aplicada a la discapacidad no verbal, con enfoque en diagnóstico del autismo y la comunicación.	<ul style="list-style-type: none">• ARASAAC, 2023.• Lampos et al., 2021.• Raja y Kannimuthu, 2023.

Tabla 1. Estructura del Estado del Arte. Elaboración propia.

Hasta la fecha, cada campo que ha sido tocado por la IA ha experimentado una revolución notable. Este fenómeno se extiende incluso al mundo académico, donde nos hemos visto notoriamente afectados por la introducción de nuevas plataformas destinadas a la creación de textos y diálogos. El desarrollo de estas herramientas educativas basadas en tecnologías de IAG, como son los modelos de lenguaje de gran tamaño (LLM), ha experimentado un crecimiento acelerado (Peñalvo et al., 2024). La aceleración de dicho desarrollo se debe en gran medida a la popularidad de modelos como ChatGPT. Sin embargo, el sistema educativo se enfrenta al reto de incorporar

estas nuevas modalidades de aprendizaje que presenta la IA, mientras se mantiene arraigado en métodos tradicionales de enseñanza y se encuentra al mismo tiempo en un contexto de abundancia de información (Peñalvo et al., 2024).

A pesar de que la IAG es una disciplina relativamente novedosa, se ha acumulado una cantidad considerable de investigaciones sobre la misma. En concreto, se ha analizado la evolución de los modelos de generación de imágenes a partir de descripciones textuales, o *text-to-image*, logrando avances significativos a través de modelos de difusión, destacando algunos como el Denoising Diffusion Probabilistic Model (DDPM) o el Stable Diffusion (Zhang et al., 2023). Dichos modelos han evidenciado su capacidad para generar imágenes de alta calidad que se ajustan fielmente a los textos proporcionados (Zhang et al., 2023). No obstante, el estudio también reconoce la falta de investigaciones exhaustivas en ciertas áreas, como la evaluación completa de estos modelos de difusión en el proceso de generación de imágenes (Zhang et al., 2023). También se han analizado funcionalidades similares, como la síntesis de imágenes a partir de texto, que posee una relevancia significativa en los medios de comunicación (Franganillo, 2023). Como describe Franganillo (2023), estos modelos adquieren patrones y rasgos de imágenes de entrenamiento para producir nuevas imágenes que reflejen dichos patrones, incluso aprovechando el contenido en línea para su aprendizaje.

Sakirin y Kusuma (2023) realizan un análisis más profundo, diferenciando las tres grandes familias algorítmicas en las que se dividen las técnicas de los modelos largos de lenguaje: Generative Adversarial Networks (GANs), Autoencoders y Autoregressive Models. Cada enfoque algorítmico contribuye de manera única a la creación de imágenes; los GANs logran la generación de imágenes realistas mediante un proceso competitivo entre dos redes; los Autoencoders crean imágenes al compactar y reconstruir datos; y los Autoregressive Models generan imágenes secuenciales. Llevan a cabo una revisión exhaustiva de diferentes enfoques, modelos y aplicaciones, además de analizar innovaciones en pérdidas, arquitecturas neuronales y técnicas de entrenamiento. Los resultados presentados comprenden una amplia gama de avances en la creación de imágenes, generación de audio, síntesis de video o elaboración de modelos 3D.

En resumen, la IAG sigue en una constante evolución, lo que indica que todavía hay un amplio margen para la realización de análisis adicionales. Los desarrollos colectivos reflejan un avance considerable en la capacidad de estos modelos fundacionales para generar contenido, aunque requieran un mayor desarrollo y trabajo de investigación.

Existe una creciente sensibilización acerca de los trastornos del lenguaje en la sociedad, además de disponer de una mayor cantidad de información al respecto (Artigas-Pallarès et al., 2022). En concreto, el trastorno del desarrollo del lenguaje emerge como uno de los trastornos del desarrollo más frecuentes y que mayor impacto tienen, ya que no solo afecta a la comunicación, sino que también repercute en otros ámbitos fundamentales como es el aprendizaje (Artigas-Pallarès et al., 2022). Las patologías del lenguaje presentan una diversidad considerable, constituyendo un amplio rango de trastornos que afectan desde la comprensión del lenguaje hasta la producción del mismo (Rupérez et al., 2022). En particular, el Trastorno del Lenguaje (TL) se caracteriza por la presencia de dificultades en la recepción e implementación del lenguaje, que no impactan únicamente en la comunicación oral, sino también en la escrita y en el lenguaje de signos (Rupérez et al., 2022).

Entre las alternativas de comunicación que poseen los colectivos con dificultades en el habla encontramos los gestos y señas, los Sistemas Alternativos y Aumentativos de Comunicación (SAAC) o los símbolos y pictogramas. Estas modalidades de comunicación desempeñan un papel crucial en la superación de los obstáculos en el habla y el lenguaje, facilitando así la comunicación efectiva y una mejor participación en interacciones sociales para estos grupos (Rupérez et al., 2022). En concreto, el presente proyecto se centrará en el espectro autista, aunque su aplicación podría extrapolarse a otros colectivos.

En lo que respecta a la combinación de la IA y la comunicación de los colectivos con discapacidad no verbal, la gran mayoría de los estudios encontrados se centran en el diagnóstico del autismo. Por ejemplo, el estudio realizado por Raja y Kannimuthu (2023) exploró métodos para procesar y clasificar datos de resonancia magnética funcional (fMRI) con el objetivo de identificar patrones asociados con el autismo. Los resultados obtenidos incluyen mejoras en la detección del autismo a partir de datos de

neuroimagen, además de la identificación de patrones específicos en los datos de fMRI, que pueden diferenciar entre individuos con autismo y controles sanos. Otro ejemplo es el análisis que reflejan Anagnostopoulou et al., (2020) sobre las técnicas de aprendizaje automático y su potencial para aumentar la exactitud y eficacia del diagnóstico al examinar extensas colecciones de datos e identificar tendencias significativas en el comportamiento y características neurobiológicas de individuos dentro del espectro autista.

Sin embargo, de la misma forma se han identificado investigaciones sobre las diferentes estrategias de comunicación. En el estudio de Lampos et al, (2021) se examina cómo las diferentes formas en que los profesores se comunican con los alumnos con Trastornos del Espectro Autista (TEA) afectan a sus respuestas en las clases. Se llevó a cabo un experimento innovador que utilizó aprendizaje automático e IA para simular las interacciones entre maestros y estudiantes en tiempo real. Los investigadores, afiliados a la Universidad College London, desarrollaron un protocolo para recopilar datos de estas interacciones, los codificaron y realizaron un análisis estadístico para identificar patrones. Los resultados destacaron el potencial de la IA y el aprendizaje automático para mejorar la práctica educativa en este contexto.

Otro estudio que cabe destacar es el realizado en colegios de Reino Unido, donde probaron la eficacia de la tecnología de intervención ECHOES (Enriching Childhoods with Opportunities in Education Software) en niños autistas de bajo rendimiento (Porayska-Pomsta et al., 2018). Esta herramienta emplea la IA con la acción humana, donde los niños interactúan con un agente virtual a través de un dispositivo electrónico con pantalla táctil. De esta manera, los alumnos se podían comunicar con los profesores sin necesidad de una interacción directa con los mismos. Los resultados finalmente mostraron un aumento significativo en las respuestas de los niños, así como tendencias positivas en sus iniciativas hacia los humanos y los agentes virtuales.

También encontramos instituciones como la de ARASAAC (2023), un Centro Aragonés dedicado a la Comunicación Aumentativa y Alternativa, donde se promueven y desarrollan los Sistemas Aumentativos y Alternativos de Comunicación (SAAC), enfoques de expresión distintos al lenguaje oral. Dichos sistemas se basan en el empleo de representaciones visuales, comúnmente conocidas como pictogramas, que simplifican la interacción a las personas que tienen dificultades en este ámbito debido a diversas razones (diversidad funcional, desconocimiento del idioma, traumatismos y

degeneración cognitiva). La relevancia del tema tratado reside principalmente en facilitar la comunicación del espectro autista verbal y no verbal, que es posible a través de diversos dispositivos electrónicos con el uso de pictogramas. El problema reside en que los pictogramas que recoge el proyecto ARASAAC para la comunicación del espectro autista son tan limitados como fijos.

Son muchas las instituciones que implantan la tecnología como mecanismo de ayuda para la comunicación de sus alumnos con discapacidad. No obstante, este marco teórico deja en evidencia que no existen investigaciones que empleen la IA para mejorar la generación de pictogramas y, en consecuencia, la comunicación del espectro autista. Únicamente se encuentran análisis sobre la mejora de la comunicación de los autistas a través de IA enfocadas a la interacción profesor – alumno. Por lo tanto, la justificación del tema tratado en este proyecto es de gran relevancia.

3. ALCANCE DE LA TESIS DE FIN DE GRADO

La siguiente sección de la tesis tiene como objetivo delimitar claramente el alcance y las condiciones bajo las cuales se llevará a cabo la investigación. En este sentido, se abordan diversos aspectos fundamentales que guiarán el desarrollo de la tesis, proporcionando una visión detallada de los objetivos, hipótesis, asunciones y restricciones que estructuran el marco de la investigación. A continuación, se presenta un listado detallado de los mismos.

Listado de objetivos. Los objetivos recogen las metas específicas que se pretenden alcanzar durante el desarrollo de la tesis.

O1: Redactar un documento para la Tesis de Fin de Grado que recoja de manera exhaustiva toda la información relevante, proporcionando así un recurso completo y detallado.

O2: Utilizar ChatGPT para la creación de pictogramas, buscando generarlos de manera instantánea y permitiendo representaciones gráficas más versátiles.

O3: Realizar un clasificador de imágenes que permita diferenciar entre los pictogramas creados por ARASAAC y aquellos creados por la IAG.

Listado de hipótesis. Las hipótesis se plantean como las afirmaciones que se pretenden verificar en relación con los datos recopilados y el análisis realizado.

H1: Los pictogramas generados por ChatGPT *Custom* son comparables a los pictogramas provenientes de ARASAAC.

Listado de asunciones. Las asunciones se definen como las premisas consideradas verdaderas para el propósito de esta tesis.

A1: Presunción de que la población objetivo tiene la capacidad de comprender los pictogramas y utilizarlos de manera efectiva como herramientas de comunicación y representación visual.

A2: Se parte de la premisa de que la mayoría de los pictogramas que generados por ChatGPT son válidos y pertinentes para su uso en el contexto de la investigación.

Listado de restricciones. Las restricciones constituyen las limitaciones del alcance y desarrollo de la tesis bajo las cuales se lleva a cabo la investigación.

R1: Imposibilidad de incorporar pictogramas del ARASAAC al modelo debido a restricciones de derechos de autor que impiden su inclusión directa en el trabajo.

R2: La Tesis de Fin de Grado está sujeta a una restricción de horas dedicadas, lo que significa que el trabajo recibirá exclusivamente la dedicación disponible dentro de ese marco temporal definido.

R3: Incapacidad para desarrollar un entrenamiento de un modelo largo de lenguaje que genera imágenes desde cero debido a ciertas limitaciones, tales como la disponibilidad de recursos informáticos y el tiempo requerido para dicho desarrollo.

4. DESCRIPCIÓN DEL MODELO

Respecto a la metodología, esta investigación se caracteriza por el empleo de modelos largos de lenguaje para la creación de imágenes. De manera inductiva, se han analizado los diferentes modelos de IA para la creación de imágenes a partir del texto recibido, llegando a la conclusión de que el modelo que más se ajusta para este proyecto es el de DALLE-3, a través de la plataforma de ChatGPT-4.

ChatGPT es un modelo de IA que simula conversaciones humanas a través del Procesamiento del Lenguaje Natural (PLN), utilizando una arquitectura de Transformadores y técnicas de aprendizaje profundo. Este modelo se entrena previamente con una amplia gama de textos de internet, aprendiendo diversos patrones de lenguaje, gramática, y conocimientos diversos, lo que le permite predecir palabras en secuencias de texto, ajustando millones de parámetros para ser capaz de afinar sus respuestas. Los Transformadores analizan el contexto de cada palabra dentro de una frase, lo que posibilita a ChatGPT generar texto, responder preguntas y realizar tareas lingüísticas específicas con alto grado de coherencia y relevancia. Su entrenamiento en vastos conjuntos de datos le otorga la capacidad de predecir la siguiente palabra en una oración basándose en el contexto previo, ayudándose de redes neuronales artificiales para simular el proceso cognitivo humano en el momento de generación de lenguaje (Carballar, 2024).

DALLE-3 es un modelo apto para este proyecto debido a su eficacia en la creación de imágenes basadas en las descripciones textuales, por lo que su adaptación a la generación de representaciones visuales dinámicas e inmediatas es idónea. El modelo se trata de la tercera versión de un modelo de generación de imágenes desarrollado por OpenAI, que mejora significativamente la generación de imágenes a partir de descripciones textuales. A diferencia de sus predecesores, DALLE-3 se entrena con descripciones de imágenes altamente detalladas, lo que le permite comprender no solo las indicaciones, sino también la historia detrás de ellas. El modelo se destaca por su potente generador de subtítulos de imágenes, el cual se alimenta de una combinación de descripciones sintéticas y reales para lograr resultados excepcionales. Este generador de subtítulos utiliza una arquitectura similar a la de un modelo de lenguaje, pero en lugar

de procesar texto, se enfoca en imágenes, convirtiéndolas en una representación comprensible para el modelo (Bouchard, 2023).

Por lo tanto, se creó un ChatGPT personalizado que, a través de la herramienta de DALLE-3, generase pictogramas a partir de descripciones textuales como "comer" o "manzana". Se configuró de la misma forma el sistema con instrucciones específicas orientadas a facilitar la comunicación visual, diseñando este enfoque especialmente para ayudar a personas con dificultades de comunicación, como individuos con autismo, a través de la creación de pictogramas intuitivamente reconocibles. También se indicaron al Chat directrices claras para que los colores característicos se asociasen con ciertos elementos; por ejemplo, un plátano sería amarillo, un balón de baloncesto, naranja, y el arroz, blanco. Al representar personas, se utilizan tonos naturales para el color de la piel, buscando una representación realista. Otra instrucción que se proporcionó al Chat customizado fue que se priorizase la simplicidad a la hora de crear los pictogramas y que no se centrara en los detalles específicos de las imágenes.

Para evaluar y ampliar la utilidad de este ChatGPT personalizado, se extrajeron 100 pictogramas generados por este sistema y los mismos 100 pictogramas de la página web de ARASAAC, la conocida fuente de recursos gráficos para facilitar la comunicación previamente mencionada. El propósito será intentar predecir a través de Redes Neuronales Convolucionales (CNN) si una imagen pertenece a ARASAAC o fue generada por DALLE-3.

Las CNN son una clase especializada de redes neuronales diseñadas para procesar datos con una estructura de cuadrícula, como imágenes. Su arquitectura típicamente incluye tres capas principales: la capa de convolución, la capa de agrupación y la capa completamente conectada (Mishra, 2021; Saha, 2018).

- **La capa de convolución** se encarga de aplicar filtros a pequeñas áreas de la imagen para detectar características locales como bordes y texturas, lo que produce un mapa de características que se utiliza para aprender patrones visuales complejos. La ecuación básica para la operación de convolución es la siguiente:

$$W_{out} = \frac{W - F + 2P}{S} + 1 \quad \text{Ecuación (1)}$$

Esta fórmula calcula el tamaño de la salida de una capa de convolución dados el tamaño de entrada W , el tamaño del filtro F , el paso S y el relleno P (Mishra, 2021).

La siguiente figura (Ilustración 1) proporciona una representación visual de la capa de convolución.

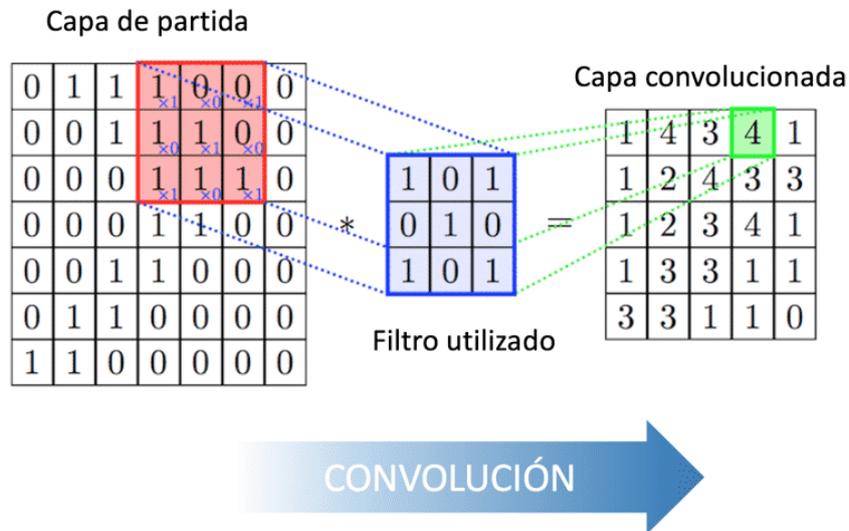


Ilustración 1. Convolución, Calvo, 2018.

- **La capa de agrupación**, como el *max-pooling*, se ocupa de reducir la dimensionalidad de la representación de la imagen al resumir la información en regiones locales. Como consecuencia, se mejora la eficiencia computacional y la invariancia a la translación en la red (Mishra, 2021; Saha, 2018).

La siguiente fórmula calcula el tamaño de salida de una capa de agrupación (*max-pooling*) dados el tamaño de entrada W , el tamaño del filtro F y el paso S (Mishra, 2021).

$$W_{out} = \frac{W - F}{S} + 1 \quad \text{Ecuación (2)}$$

- **La capa completamente conectada** finalmente transforma la salida de las capas anteriores en una forma adecuada para la clasificación o regresión. Dicha transformación permite capturar relaciones complejas entre las características extraídas y realizar tareas como reconocimiento de objetos y clasificación de

imágenes (Mishra, 2021; Saha, 2018). La imagen que se encuentra a continuación presenta una representación visual del funcionamiento de una red convolucional completa, mostrando también la capa completamente conectada.

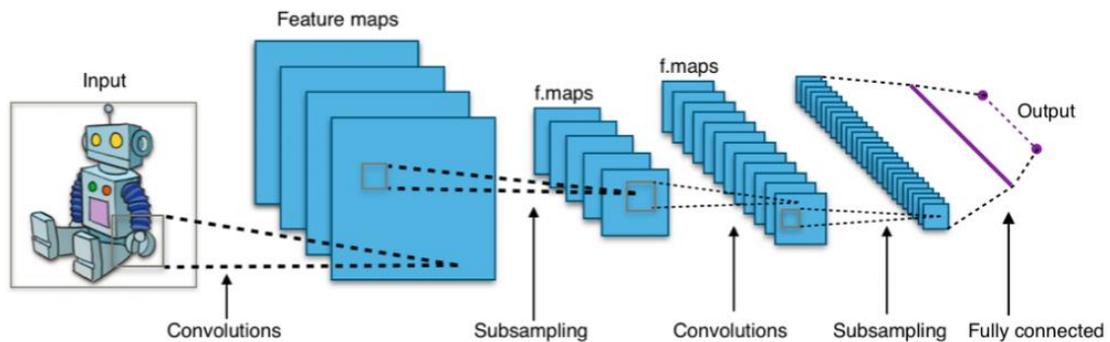


Ilustración 2. Red Convolucional, Costa, 2020.

Las CNN aprenden representaciones jerárquicas de los datos, comenzando con características simples en las capas iniciales y progresando hacia características más abstractas en las capas posteriores. Además, han demostrado ser efectivas en una variedad de tareas de visión por computadora, como reconocimiento de objetos y segmentación de imágenes, lo que las convierte en elementos fundamentales en el campo del aprendizaje profundo y la IA (Mishra, 2021; Saha, 2018).

5. DESCRIPCIÓN DEL PROBLEMA

A lo largo de los años, se han explorado diversas estrategias y enfoques con el fin de facilitar una comunicación no verbal eficaz en miembros del espectro autista. Por lo tanto, se ha creado un amplio rango de métodos y técnicas diseñadas específicamente para abordar las necesidades particulares de este colectivo.

Por ejemplo, existen terapias como la terapia ABA (Análisis Conductual Aplicado) y la terapia TEACCH (Tratamiento y Educación de Niños con Autismo y Problemas de Comunicación Relacionados) que tratan de enseñar diversas habilidades de comunicación no verbal a las personas en el espectro autista. De un lado, la terapia ABA se basa en el aprendizaje estructurado de las habilidades, dividiéndolas en pequeñas partes y enseñándolas de una en una, con un inicio y un final delimitados por estímulos o reforzados. Y, de otro, la terapia TEACCH intenta estructurar en un contexto más social y educativo, priorizando el empleo de estímulos concretos sobre los refuerzos, empleando técnicas como el uso de ambientes predecibles, incluso con apoyo visual. (Guzmán et al., 2017).

Asimismo, en Guzmán et al. (2017) también se aborda el concepto de "Proceso de comunicación mediado por asistencia" como una forma de establecer un puente entre personas con TEA y su entorno. La idea se enfoca en estrategias que utilizan sistemas de comunicación aumentada y alternativa (AAC) en ordenadores, móviles o tabletas con software que ofrece múltiples funciones y una rápida acción – reacción.

Un ejemplo es el Sistema de Comunicación por Intercambio de Imágenes (PECS), que utiliza pictogramas para construir una vía de comunicación progresiva. Sin embargo, existen barreras logísticas en su implementación, como puede ser el deterioro de los materiales físicos utilizados, que podrían mejorarse con el uso de tecnología. Además del PECS, otras tecnologías asistenciales como la robótica y los entornos virtuales también están siendo exploradas para su aplicación en el TEA. Los robots y entornos virtuales ofrecen interacciones predecibles y repetitivas, lo que puede ser preferido por algunos individuos con TEA, y proporcionan una forma controlada de vínculo social. Los entornos de realidad virtual, por otro lado, permiten presentar tareas de aprendizaje de manera realista pero dentro de un entorno simplificado, lo que facilita la

participación del individuo en actividades terapéuticas y educativas (Guzmán et al., 2017).

Pinel y Rendón (2018) discuten más a fondo los robots sociales enfocados a facilitar la interacción social a través de juegos o lecciones educativas. Al ser juguetes interactivos, son capaces de comunicarse y establecer relaciones sociales. Es por ello por lo que han resultado beneficiosos y de gran apoyo en terapias para personas con TEA, mejorando la comunicación y reduciendo comportamientos repetitivos. Aunque presentan un gran potencial, su coste y la falta de evidencia a largo plazo sobre su efectividad son desafíos importantes. Además de la necesidad de supervisión humana para garantizar resultados positivos en las sesiones terapéuticas.

Desde otro enfoque encontramos la revisión de las diversas intervenciones destinadas a mejorar la comunicación social en niños con TEA de Baixauli-Fortea et al. (2017). Entre las estrategias mencionadas se encuentran las historias sociales, que son narrativas breves diseñadas para ayudar a los niños a comprender mejor el mundo social que les rodea. Las historias se apoyan de situaciones sociales relevantes, utilizando un lenguaje adaptado al nivel del niño y centrándose en sus preferencias por el procesamiento visual. Otro método son las conversaciones en viñetas, una técnica propuesta por Carol Gray, en la que los niños participan activamente en la elaboración de una historia visual que aborda situaciones sociales problemáticas y busca identificar soluciones. Además, se destaca el video modelaje como un método de instrucción que utiliza videos para enseñar habilidades sociales. Los niños con TEA pueden aprender observando a otros ejecutar comportamientos específicos, lo que les permite comprender mejor las señales no verbales y aprender a responder adecuadamente en diversas situaciones sociales. La cuarta y última estrategia se constituye por las intervenciones mediadas por pares, que implican enseñar a los compañeros del niño autista posibles estrategias para facilitar la comunicación social y responder a comportamientos específicos. En este caso, se busca primordialmente generar un clima más comprensivo y solidario, donde los compañeros puedan brindar apoyo y comprensión, aunque la edad pueda presentar dificultades a la hora de comprender la importancia de este tema.

En definitiva, a través de numerosas estrategias y tecnologías se ha buscado mejorar la comunicación no verbal en individuos del espectro autista. Desde terapias estructuradas

como ABA y TEACCH, hasta tecnologías innovadoras como el PECS y la robótica social, además de muchas otras. Por lo tanto, el desarrollo de estas herramientas prácticas y centradas en el individuo para abordar las necesidades específicas de comunicación en el ámbito social, sigue en continuo movimiento.

En este sentido, la motivación subyacente del presente trabajo es determinar si los pictogramas generados por la IA son indistinguibles de los pictogramas de ARASAAC. Es decir, si la IA puede llegar a alcanzar un nivel de realismo y de efectividad que haga que sean muy difíciles de diferenciar de los pictogramas reales. Dicho análisis no solo se centra en los resultados finales, sino en el proceso de evaluación en sí mismo. Al profundizar en la metodología utilizada para contrastar los pictogramas generados con los reales, se busca comprender mejor los límites y las posibilidades de la tecnología actual en este campo específico. Además, llevar a cabo esta investigación no solo arroja luz sobre la efectividad de los pictogramas generados por IA, sino que también destaca áreas clave para futuras mejoras y desarrollos en la generación de contenido visual.

Para ello, se emplea un conjunto de datos que consta de 100 imágenes de DALLE y 100 imágenes de ARASAAC. Se dividen en un 70% para entrenamiento, un 10% para validación y un 20% para pruebas.

La arquitectura de la red neuronal convolucional utilizada sigue un diseño típico para problemas de clasificación de imágenes. Comienza con una capa de convolución que emplea 32 filtros de tamaño 3x3 y utiliza la función de activación ReLU para introducir no linealidad en la red. La capa espera imágenes de entrada con dimensiones de 150x150 píxeles y 3 canales RGB. Después de cada capa de convolución, se aplica una capa de Max Pooling con una ventana de pooling de tamaño 2x2 para reducir la dimensionalidad y extraer características importantes. Este proceso se repite dos veces más, aumentando el número de filtros a medida que se avanza en la red para capturar características más complejas.

Después de las capas convolucionales y de pooling, se aplica una capa de aplanado para convertir la salida en un vector unidimensional. Esto es esencial para conectar las capas convolucionales con las capas densamente conectadas que siguen. Luego, se añaden dos

capas densamente conectadas. La primera tiene 512 neuronas y utiliza ReLU como función de activación. La última capa densa tiene una sola neurona y utiliza la función de activación sigmoid, comúnmente empleada en problemas de clasificación binaria para producir una salida en el rango [0, 1]. Esta arquitectura está diseñada para extraer características relevantes de las imágenes de entrada y clasificarlas de manera efectiva en dos categorías distintas. El código descrito se encuentra en el siguiente repositorio: https://github.com/EduardoGarrido90/tutorials_data_science/blob/main/cnn_class_ara_i_ag.py

La red se entrena utilizando el optimizador Adam para maximizar la precisión. Además, se realiza un preprocesamiento de los datos mediante un reescalado, lo que implica ajustar los valores de los píxeles de las imágenes para que estén en un rango específico, para así mejorar la velocidad y la estabilidad del proceso de entrenamiento.

6. EXPERIMENTOS Y RESULTADOS

Tras extraer las 200 imágenes, se procedió a homogeneizarlas para unificar su formato. Como las imágenes de DALLE-3 estaban en formato webp, se convirtieron a png para que coincidieran con las de ARASAAC. Además, se modificó el fondo transparente de las imágenes de DALLE-3 por un fondo blanco, ya que de lo contrario la red neuronal convolucional podría clasificar las imágenes en función del color de fondo en lugar del contenido principal de la imagen.

Seguidamente, se llevó a cabo una reducción de calidad en todas las imágenes para igualar el número de píxeles y mantener un tamaño adecuado para la red neuronal convolucional, lo que ayuda a reducir el costo del entrenamiento. Esta decisión se tomó teniendo en cuenta que el problema en cuestión era el reconocimiento de pictogramas simples, por lo que no se necesitaba una alta resolución en las imágenes.

El siguiente paso fue dividir las imágenes en un 80% para entrenamiento y un 20% para evaluación del modelo. Aunque esta división puede estar sesgada, dada la diversidad de pictogramas, se consideró suficiente para los propósitos del proyecto. Es cierto que para alcanzar una mayor robustez se podría emplear un método de evaluación del error predictivo más fiable, como un *k-fold cross-validation*, pero para este caso se consideró que esta división era la adecuada.

En cuanto a los resultados, el conjunto de *batch* era de 20 y se entrenó durante 10 épocas. Las métricas de precisión y pérdida obtenidas en el conjunto de entrenamiento y validación son las siguientes:

- Precisión de entrenamiento (accuracy): 0.9931
- Pérdida de entrenamiento (loss): 0.0401
- Precisión de validación (val_accuracy): 0.9500
- Pérdida de validación (val_loss): 0.0976

Si bien la red convolucional alcanza altos niveles de precisión en el conjunto de entrenamiento y validación, logrando una precisión de entrenamiento del 99.31% y una precisión de validación del 95%, la verdadera prueba de la efectividad de los

pictogramas generados por IA radica en su capacidad para superar el conjunto de prueba.

La red neuronal alcanzó un 88% de precisión en el conjunto de test (20% de las imágenes). Por lo tanto, se puede concluir que ambos conjuntos de imágenes aún son radicalmente distintos y se nota que los generadores no imitan bien el estilo. Primordialmente, este resultado se debe a que son únicamente 100 imágenes repartidas en *train*, validación y *test*, por lo que una red convolucional muy simple con un preprocesamiento básico ha sido capaz de clasificarlas de manera óptima. En definitiva, los hallazgos sugieren que los pictogramas de ARASAAC siguen siendo claramente distinguibles y probablemente más efectivos. Aunque es valioso contar con herramientas para analizar nuevos pictogramas en una nueva manera de comunicarse, este análisis evidencia la necesidad de seguir mejorando la generación de pictogramas a través de la IA.

7. COMERCIALIZACIÓN DE LA SOLUCIÓN

La posible comercialización de esta solución sería la creación de una aplicación para tabletas y comunicadores que serviría como un puente entre la complejidad del modelo de IA y la accesibilidad para los usuarios finales. Al diseñar la interfaz de usuario, se debe tener en cuenta la diversidad de los perfiles de usuarios, incluidos aquellos con discapacidades cognitivas o físicas, garantizando así que la aplicación sea verdaderamente inclusiva. La implementación de un proceso de carga del modelo de IA sencillo y rápido, así como la integración de una función de entrada de *prompts* intuitiva, son aspectos cruciales que deben abordarse durante el desarrollo de la aplicación. Además, se deben realizar pruebas exhaustivas para garantizar que la aplicación funcione sin problemas en una variedad de dispositivos y sistemas operativos, lo que contribuirá a una experiencia de usuario óptima.

En cuanto al perfil de los consumidores, es fundamental comprender las necesidades y los comportamientos de los usuarios potenciales para diseñar una estrategia de comercialización efectiva y asegurar la adopción exitosa de la aplicación. La audiencia objetivo de esta solución abarcaría una amplia gama de usuarios, desde aficionados creativos hasta profesionales de diversos campos como el arte, el diseño, la educación y el marketing. Los aficionados creativos podrían utilizar la aplicación para explorar su imaginación y expresar su creatividad de nuevas maneras, mientras que los profesionales encontrarían en ella una herramienta poderosa para generar contenido visual de manera rápida y eficiente, lo que les permitiría ahorrar tiempo y recursos en sus proyectos.

El diagrama de flujo que se muestra a continuación detalla el proceso de desarrollo y comercialización de una aplicación diseñada para tabletas y comunicadores, destinada a facilitar la comunicación de personas con TEA mediante el uso de pictogramas generados por un modelo de IA.

El primer recuadro de inicio marca el comienzo del proceso de desarrollo y lanzamiento de la aplicación. Seguidamente, se finalizaría el desarrollo y refinamiento del modelo de generación de pictogramas, asegurando su eficacia y precisión en la interpretación de las expresiones comunicativas.

El siguiente paso sería crear una aplicación básica con la capacidad de cargar el modelo, estableciendo una estructura funcional que permita la incorporación y el uso eficiente de este modelo dentro de la aplicación. Por lo tanto, también implica la creación de una interfaz simple que facilite la carga del modelo y la preparación para su uso posterior en la generación de imágenes. Una vez completada, el modelo se integra en la aplicación de comunicación, donde los usuarios pueden acceder a él de manera intuitiva y sencilla.

Por otro lado, la creación de una interfaz de usuario accesible es una etapa fundamental del proceso, considerando la diversidad de perfiles de usuarios, incluidos aquellos con discapacidades cognitivas o físicas. La interfaz se diseña con el objetivo de garantizar la inclusión y la facilidad de uso para todos los usuarios, independientemente de sus habilidades o limitaciones y con un especial énfasis en miembros del colectivo autista.

Después, se llevarían a cabo unas pruebas exhaustivas para evaluar la funcionalidad y la accesibilidad de la aplicación, asegurando que cumpla con los estándares de calidad requeridos. En otras palabras, garantizar que la aplicación funcione correctamente y sea accesible para el público objetivo.

Posteriormente, se optimiza la aplicación para garantizar su compatibilidad y rendimiento óptimo en una variedad de dispositivos y sistemas operativos, incluidas tabletas y comunicadores. Dicha adaptación se realiza teniendo en cuenta que muchas personas dentro del colectivo autista ya están familiarizadas y acostumbradas a utilizar este tipo de dispositivos en su vida diaria. La elección de tablets y comunicadores se basa en observaciones previas realizadas durante la investigación, donde se ha constatado su amplia utilización por personas con TEA o con discapacidades en el habla. Al hacer que la aplicación sea compatible con este tipo de dispositivos, se facilita su accesibilidad y uso para los usuarios objetivo, ya que pueden aprovechar su experiencia previa para interactuar de manera más intuitiva y efectiva con la aplicación.

El siguiente paso sería configurar el proceso de generación de imágenes dentro de la aplicación. En detalle, la aplicación procedería a cargar el modelo de creación de pictogramas, seguido de la recepción de un *prompt* por parte del usuario. El *prompt* puede variar desde una breve descripción hasta una instrucción detallada. Acto seguido, utilizando el modelo previamente cargado, la aplicación procedería a generar la imagen, cuyo contenido se basaría en la información proporcionada por el usuario.

Una vez que la aplicación ya sea compatible con los diferentes dispositivos, se llevarán a cabo eventos de lanzamiento y demostraciones en centros especializados, escuelas y organizaciones relacionadas con el TEA. El propósito de estos eventos sería mostrar con detalle cómo funciona la aplicación y cómo puede integrarse en la vida diaria de los individuos con autismo, permitiendo a profesionales y usuarios potenciales comprender a fondo su utilidad y beneficios. La retroalimentación recopilada durante estas demostraciones es esencial para identificar áreas de mejora antes del lanzamiento oficial de la aplicación.

La promoción de la aplicación se lleva a cabo a través de asociaciones y grupos de apoyo, asegurando que llegue a su público objetivo de manera efectiva. Además, se podrían aprovechar las redes existentes para aumentar su visibilidad y adopción, así como apoyarse en asociaciones como ARASAAC para dar a conocer la aplicación.

Finalmente, la aplicación se lanza oficialmente y se pone a disposición del público para su descarga y uso, añadiendo en la misma una descripción clara de sus características y beneficios para usuarios con TEA. A medida que la aplicación se utiliza y recibe retroalimentación de los usuarios, se realizan revisiones y mejoras continuas para garantizar su eficacia y utilidad a largo plazo.

La ilustración del diagrama de flujo al completo se muestra a continuación.

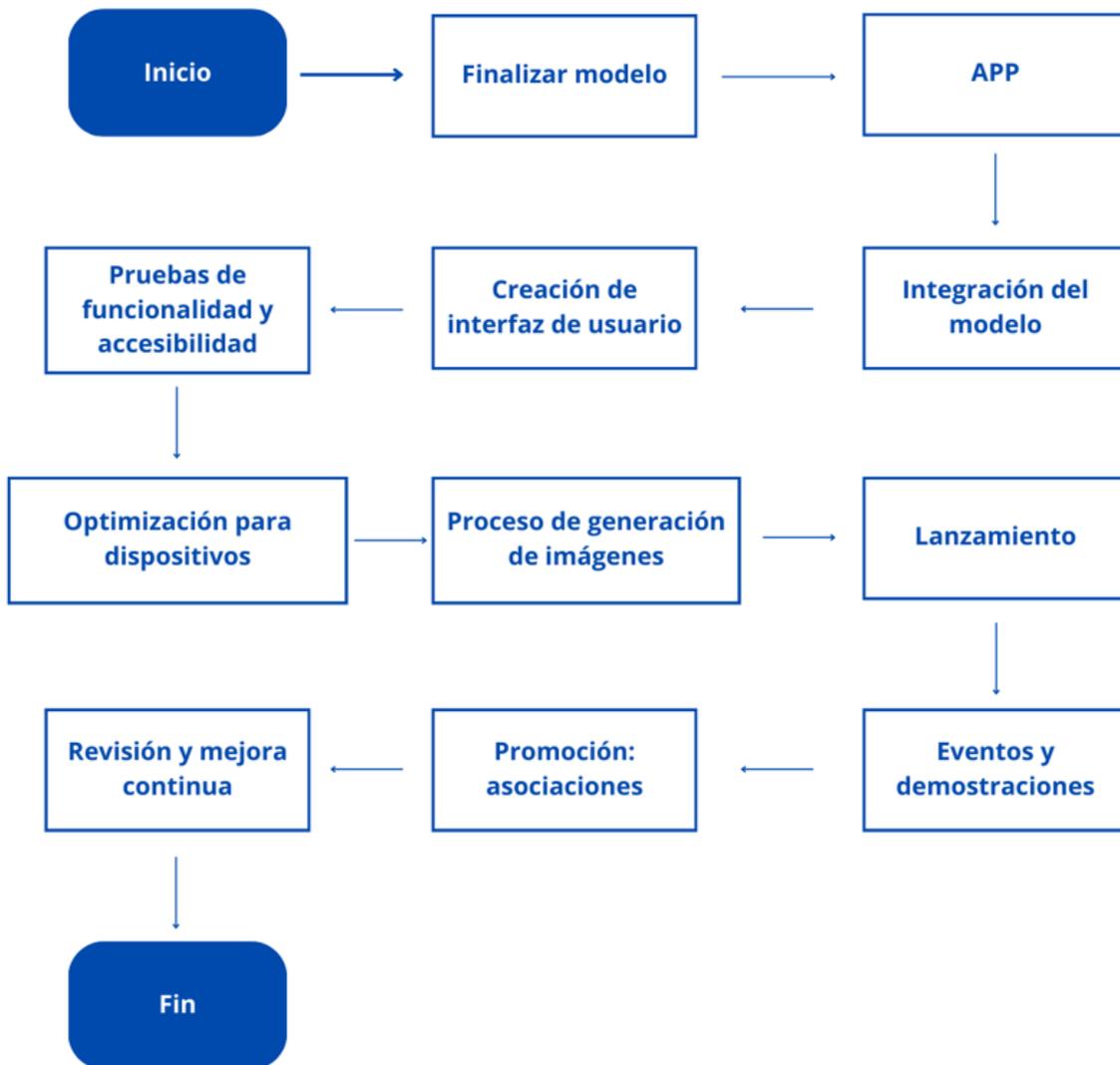


Ilustración 3. Diagrama de flujo. Elaboración propia.

8. CONCLUSIONES

Una de las primeras evidencias que se obtienen en este proyecto de investigación es el potencial de la IAG para la creación de pictogramas destinados a mejorar la comunicación de personas con dificultades verbales, como los miembros del espectro autista.

Durante la investigación, se ha observado también una evolución significativa en el campo de la IA, especialmente en la aplicación de modelos de LLM como ChatGPT en diversos contextos, incluida la comunicación para personas con discapacidad del habla.

Los resultados obtenidos sugieren que los pictogramas generados por la IA son comparables en calidad y utilidad a los proporcionados por fuentes convencionales como ARASAAC. En consecuencia, nuevas perspectivas emergen en el ámbito de la comunicación asistida, ofreciendo una alternativa innovadora y más eficiente para la creación de recursos visuales de apoyo a la comunicación. Sin embargo, continúan siendo diferenciables entre ellos, por lo que la IA aun no es capaz de imitar al completo los pictogramas que ya existen, sean creados por ARASAAC o por cualquier otra organización.

La implementación de esta tecnología no solo permite la generación instantánea de pictogramas a partir de descripciones textuales, sino que también mejora la accesibilidad al proporcionar una gama más amplia y dinámica de símbolos para la comunicación. Los avances en esta línea de desarrollo, como aquellos mencionados en el estado del arte, representan un paso significativo hacia la inclusión y la igualdad de oportunidades para las personas con dificultades de habla, al tiempo que ofrece una herramienta versátil y adaptable para facilitar la comunicación en una variedad de contextos.

Asimismo, se ha establecido un marco metodológico para la evaluación de pictogramas generados por IA, que incluye la comparación con pictogramas de referencia y el uso de redes neuronales convolucionales para la clasificación. Personalmente, considero que este enfoque metodológico puede servir como punto de partida para futuras investigaciones en el campo de la generación de imágenes asistido por la IA,

proporcionando una base sólida para la evaluación y mejora continua de estas herramientas.

En resumen, este proyecto ha tratado de investigar la combinación de la IA en el área de la comunicación asistida, demostrando el potencial de la IA para mejorar la calidad y la accesibilidad de las herramientas de comunicación para personas con dificultades verbales. No obstante, podemos concluir que la IA sigue teniendo que desarrollarse para poder alcanzar su máximo potencial en este ámbito. Aunque los pictogramas generados ofrecen una nueva manera de comunicarse, es evidente que se requiere un refinamiento adicional para lograr una mayor similitud con los originales y, por ende, una mayor efectividad en la comunicación asistida.

9. FUTURAS LÍNEAS DE INVESTIGACIÓN

Entre las posibles futuras líneas de investigación, encontramos las siguientes:

En primer lugar, la exploración de un mayor número de modelos de generación de imágenes. Es decir, experimentar con modelos más avanzados, ajustar los diferentes hiperparámetros del modelo, o explorar enfoques de aprendizaje semi-supervisado para intentar mejorar la calidad y la diversidad de los pictogramas generados. Asimismo, el desarrollo de herramientas de modelos fundacionales más específicas y especializadas es una perspectiva estimulante. A medida que la tecnología avanza, la IA lo hace también, pudiendo incorporar características como la personalización adaptativa a las necesidades individuales de los usuarios o la mejora de la interacción con el usuario. Dichas investigaciones podrían explorar cómo los distintos modelos largos de lenguaje pueden adaptarse y optimizarse para satisfacer las necesidades únicas de segmentos de clientes y contextos de uso. Así como analizar la posible personalización de pictogramas para usuarios individuales, es decir, adaptar los pictogramas generados para satisfacer las necesidades y preferencias individuales de los usuarios. Podría incluir el desarrollo de sistemas que permitan a los usuarios personalizar los pictogramas generados, como seleccionar colores o modificar detalles específicos.

Otra posible línea de investigación sería la profundización en el impacto de los pictogramas en la inclusión social de colectivos con discapacidad no verbal. Sería interesante estudiar cómo el uso de pictogramas generados por IA en la comunicación afecta la inclusión social y la participación de las personas con discapacidades del habla en diversos entornos, como la escuela, el trabajo o la comunidad. Además, la evaluación del impacto en la comunicación real es crucial. Es necesario evaluar cómo el uso de pictogramas generados por estos modelos largos de lenguaje afecta realmente a la comunicación y la comprensión de los usuarios objetivo en situaciones de la vida cotidiana. Las futuras investigaciones podrían incluir estudios de caso y ensayos clínicos para analizar la eficacia y la aceptación de estas herramientas en situaciones prácticas, así como para identificar posibles áreas de mejora.

Como se ha observado en la revisión de la literatura, cada vez existen más herramientas que emplean IA para el apoyo a profesionales de la educación o cuidadores de personas con discapacidad, especialmente en lo que respecta a la interacción o comunicación con estos colectivos. En consecuencia, este enfoque en desarrollo podría abrir otro campo de análisis en el que se investigaría el progreso de herramientas o aplicaciones basadas en IA, su efectividad, la identificación de posibles patrones o características específicas que las distinguen, así como la detección de posibles defectos o fallos para evitar su repetición en futuras implementaciones.

Finalmente, además de los pictogramas, la IA podría aplicarse en la generación de otros recursos de comunicación asistida, como sistemas de signos y gestos, para atender a una variedad más amplia de necesidades comunicativas de personas con dificultades verbales. Asimismo, es fundamental abordar consideraciones éticas relacionadas con la privacidad, la equidad y la inclusión a medida que se desarrollan y aplican estas tecnologías. Futuras líneas deberían incluir un enfoque proactivo en la incorporación de principios de diseño inclusivo y participativo para garantizar que estas herramientas sean accesibles y beneficiosas para todos los usuarios.

10. REFERENCIAS BIBLIOGRÁFICAS

- Abeliuk, A., & Gutiérrez, C. (2021). Historia y evolución de la inteligencia artificial. *Revista Bits de Ciencia*, (21), 14-21.
- Amador Hidalgo, L. (1996). *Inteligencia artificial y sistemas expertos*. Universidad de Córdoba, Servicio de Publicaciones.
- Anagnostopoulou, P., Alexandropoulou, V., Lorentzou, G., Lykothanasi, A., Ntaountaki, P., & Drigas, A. (2020). Artificial intelligence in autism assessment. *International Journal of Emerging Technologies in Learning (iJET)*, 15(6), 95-107.
- ARASAAC. (2023). Arasaac.org. <https://arasaac.org/aac/>. Último acceso el 22 de abril de 2024.
- Artigas-Pallarès, J., Pérez, I. P., & Mallofré, E. V. (2022). Trastornos del lenguaje. *Pediatría Integral*, 7, 12.
- Baixauli-Fortea, I., Roselló-Miranda, B., Berenguer-Forner, C., Colomer-Diago, C., & Grau-Sevilla, M. D. (2017). Intervenciones para promover la comunicación social en niños con trastornos del espectro autista. *Revista de Neurología*, 64(1), 39-44.
- Bouchard, L. (2023). *DALLE 3 Explained: Improving Image Generation with Better Captions*. Louis-François Bouchard, Aka What's AI. <https://www.louisbouchard.ai/dalle-3/>. Último acceso el 22 de abril de 2024.
- Calvo, D. (2018, 8 diciembre). *Red neuronal convolucional CNN*. Diego Calvo. <https://www.diegocalvo.es/red-neuronal-convolucional/>. Último acceso el 22 de abril de 2024.
- Carballar, J. (2024, 1 marzo). *Cómo funciona ChatGPT por dentro*. Carballar.com. <https://carballar.com/como-funciona-chatgpt-por-dentro>. Último acceso el 22 de abril de 2024.
- Costa, P. (2020, 1 octubre). *Redes neuronales convolucionales explicadas / POCHO COSTA*. POCHO COSTA. <https://pochocosta.com/podcast/redes-neuronales-convolucionales-explicadas/>. Último acceso el 22 de abril de 2024.

- Franganillo, J. (2023). La inteligencia artificial generativa y su impacto en la creación de contenidos mediáticos. *methaodos. revista de ciencias sociales*, 11(2), 15.
- Guzmán, G. B., Putrino, N., Martínez, F. J. R., & Quiroz, N. (2017). Nuevas tecnologías: Puentes de comunicación en el trastorno del espectro autista (TEA). *Terapia Psicológica*, 35(3), 247-258. <https://doi.org/10.4067/s0718-48082017000300247>. Último acceso el 22 de abril de 2024.
- Lamos, V., Mintz, J., & Qu, X. (2021). An artificial intelligence approach for selecting effective teacher communication strategies in autism education. *npj Science of Learning*, 6(1), 25.
- Lombana Bermudez, A. (2018). La evolución de las brechas digitales y el auge de la Inteligencia Artificial (IA). *Revista Mexicana de Bachillerato a Distancia*, 10(20), 17-25.
- Manene, L. M. (2011). Los diagramas de flujo: su definición, objetivo, ventajas, elaboración, fases, reglas y ejemplos de aplicaciones. *Recopilado el*, 22, 09-18.
- Mishra, M. (2021). Convolutional neural networks, explained - towards data science. *Medium*. <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>. Último acceso el 22 de abril de 2024.
- Peñalvo, F. J. G., Largo, F. L., & García, J. V. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. *RIED: revista iberoamericana de educación a distancia*, 27(1), 9-39.
- Pinel, V., & Rendón, L. A. (2018). Los robots sociales como promotores de la comunicación en los Trastornos del Espectro Autista (TEA). *Letras de Hoye*, 53(1), 39. <https://doi.org/10.15448/1984-7726.2018.1.28920>. Último acceso el 22 de abril de 2024.
- Porayska-Pomsta, K., Alcorn, A. M., Avramides, K., Beale, S., Bernardini, S., Foster, M. E., ... & Smith, T. J. (2018). Blending human and artificial intelligence to support autistic children's social communication skills. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 25(6), 1-35.

- Raja, K. C., & Kannimuthu, S. (2023). Conditional Generative Adversarial Network Approach for Autism Prediction. *Computer Systems Science & Engineering*, 44(1).
- Rupérez, A., Ramos, I., Machado, I., Fernández, D., Gortázar, M., & Aguilera, S. (2022). Trastornos del lenguaje, del habla y de la comunicación. Conceptos, clasificación y clínica. *Protoc dign ter pediatr*, 1(1), 19-30.
- Saha, S. (2018). *A Guide to Convolutional Neural Networks — the ELI5 way* / Saturn Cloud Blog. <https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/>. Último acceso el 22 de abril de 2024.
- Sakirin, T., & Kusuma, S. (2023). A Survey of Generative Artificial Intelligence Techniques. *Babylonian Journal of Artificial Intelligence*, 2023, 10-14.
- Zhang, C., Zhang, C., Zhang, M., & Kweon, I. S. (2023). Text-to-image diffusion model in generative ai: A survey. *arXiv preprint arXiv:2303.07909*.

Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado

ADVERTENCIA: Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, Clara Maldonado Gilarranz, estudiante de Administración y Dirección de Empresas y Business Analytics de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado "GENERACIÓN DE PICTOGRAMAS CON INTELIGENCIA ARTIFICIAL GENERATIVA PARA EL ESPECTRO AUTISTA", declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación:

1. **Traductor:** Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: 24 de abril

Firma: Clara Maldonado