

PROCEEDINGS OF LONDON INTERNATIONAL CONFERENCES

eISSN 2977-1870

Comparative Recognition Analysis of Image Accuracy: A Study of OpenAI and Gemini in Matching Original Visuals

Batyr Omurkulov*^{id}

Dovran Ylhamov**^{id}

Merve Kevser Gokgol***^{id}

Abstract

It is critical for developers and users to comprehend the capabilities of various models in the quickly developing field of image generation. By comparing the image output quality of Gemini and OpenAI, this study seeks to shed light on their relative advantages. We can learn more about how each model interprets visual data by concentrating on features like object recognition, color fidelity, and detail resolution. In the end, this study adds to the larger conversation about the efficiency of machine learning methods in producing superior visual content. This study compares the quality of images generated by Gemini and Open AI, focusing on how closely these images match the original ones. Both Gemini and OpenAI use advanced machine learning models to transform visual images, but they do so using different methods and data. The comparison examines key factors such as color accuracy, detail, and how well each model recognizes objects and scenes. The results show that Gemini tends to produce images with more accurate colors and fine details, while Open AI is better at understanding and interpreting a wider range of descriptions. This highlights the strengths and weaknesses of each model in producing high-quality images.

Keywords: Image Generation, Gemini, OpenAI, Color Accuracy, Image Detail



<https://doi.org/10.31039/plic.2024.12.282>

*HS of Endeavor at Austin, USA, 06.batyr@gmail.com

** HS of Endeavor at Austin, USA, 08.dovran@gmail.com

*** HS of Endeavor at Austin, USA, mervekevser@gmail.com

14th London International Conference, October 5-7, 2024



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

1. Introduction

Image recognition is one of the most important uses of artificial intelligence (AI), allowing machines to gather and interpret visual data from their surroundings. Devices, faces, messages or other visual representations are interpreted in this era. Due to the widespread use of the image from security infrastructure, digital images were created. Since self-sufficient motors, many industries have undergone revolutionary changes. Health care and electronic business. OpenAI is an organisation that is leading some of the best AI innovations. OpenAI is often known for its language processing and knowledge acquisition models. It has achieved great success in terms of information and response to human language. However, it also does additional important work on the popularity of images. OpenAI's tools can process image recordings and derive meaning from those recordings.

In the last decade or so, the field of artificial intelligence (AI) has seen remarkable advances, and Chat Generative Pre-trained Transformer is a striking example of this progress (D. Jason Slone & McCorkle, 2019). Gemini is another advanced AI tool that focuses on image analysis and popularity. This system can discover, validate, and create a complete interpretation of device and context factors with greater accuracy. Technologies like OpenAI and Gemini are pushing the boundaries of image recognition. Make a package for health care from evaluating clinical images surveillance structure to the use of conservation cameras.

Evaluating these powerful features, OpenAI prides itself on its extensive language processing capabilities coupled with its photo reputation capabilities, while Mithuni focuses specifically on photo analytics.

2. Literature Review

The power of AI aside, we are more interested in knowing the circumstances under which human analysts retain their advantage (Crawford, 2021/2021). AI was invented in 1956 by John McCarthy, but did not become as impactful till the 1970s (Ford, 2018). It is a broad field that includes machine learning and deep learning technologies (Ford, 2018). Artificial intelligence is used in many areas. Including healthcare for disease analysis and medical construction plans. Education for personal learning, autonomous cars, automotive organisations, optimising investments, fraud tracking and financial detection, customer support via chatbot, and is widely used in agriculture.

We select a set of state-of-the-art machine learning models and build our AI analyst based on an ensemble model. Our AI analyst is able to beat human analysts as a whole (Blitzstein & Hwang, 2015). Image recognition is used in many areas of life, including security structures for authentication through facial recognition, medical snapshots for diagnosis, and healthcare. It is widely used in businesses for example AI analysts outperform in 54.5% of the stock return predictions made by all I/B/E/S analysts during the sample period of 2001 to 2018 (Marr, 2015). Automatic product recognition and automatic engines for environmental sensing. And to tag images and movies on social media platforms. Also, normal life has become easier thanks to eras such as virtual assistants. And smart home appliances that understand customers play an important role among them. Continuous



developments in artificial intelligence (AI) technologies are expected to bring innovations to the future of healthcare (Allen, 2019).

3. Methodology

A comparative evaluation was used to evaluate the performance of two superior visualisation models, Gemini and OpenAI, in terms of output quality. The assessment makes a speciality of 3 major factors: shade accuracy, photograph details, and item popularity. Both models paint to generate new pictures primarily based on these inputs. OpenAI’s model is primarily based on a neural network structure that interprets text descriptions and generates photos. Gemini specialises in extracting detail and precision in pictures. Both models are pre-trained on an intensive dataset. However, Gemini’s dataset focuses more on colour matching and high-resolution info. They were blended with diverse notifications. The generated snap shots are evaluated using 3 main criteria: coloration accuracy, which measures how intently the colours of the generated photograph shape the original picture; photograph information, which evaluates the readability and sharpness of features, which include textures and edges. The capacity to correctly pick out models, items, and scenes. Human evaluators, inclusive of laptop imaginative and prescient experts for layout, will take a look at the snapshot facet with the aid of aspect with the original drawings.

I. Input Prompt for Sample 1: Draw for me The Casio Gray MTP VDO1D watch grey metal case with black background.

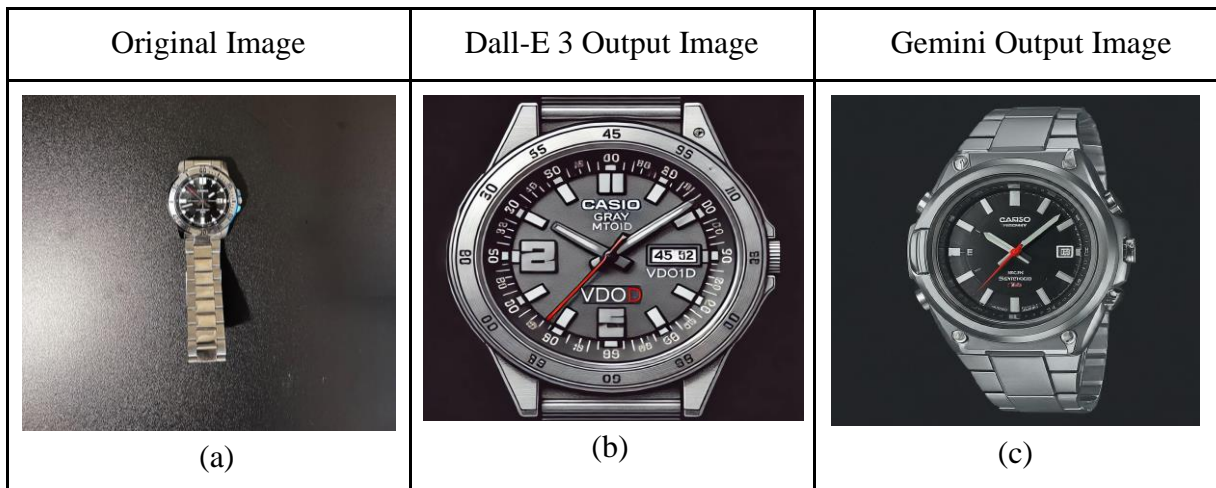


Image1. Comparison of Image 1. (a) from both AI generation tools

Input Prompt #1 The first image is the real object. The second image is the image generated by Open AI(Dall-E 3). The third image is the image generated by Gemini. Now tell the differences between them specifically and tell their similarities in percentage. Which one is similar to the real object and what percentage?

Result and Analysis: The Open AI-generated image has a higher similarity rate when compared to Gemini generated image. Similarity is mainly in overall style and the presence of the Casio brand name. However, the image generated by both (Gemini, and Open AI) contains significant differences in details and complexity compared to the simpler design of the real watch.

Differences between Gemini and Dall- E Image Outputs: As could be seen in the Image 1 while the real watch brand is Casio, while the AI-generated images from Dall-E 3 and Gemini show "CASIO" and "CARISO" respectively. The real watch does not show specific model information on its dial, while the AI images show "GRAY MTOID" and "TWISCRA4V" respectively. The real watch has a simple date window, the Open AI image shows "45 32", while the Gemini image does not show a clear date. The real watch has a simpler dial design compared to the more complex designs in the AI images. The AI images show more obvious and detailed bezel markings than the real watch.

Similarities of Gemini and Dall E Image Output: Both Outputs were compared in percentage to analyse the similarity with the original image using [Claude-AI](#) (Claude, 2024) Both images show stainless steel watches with dark dials as in the original sample. Overall shapes and analog dials and red coloured second hands are also common features for all generated images. Similarity percentage is given in Table 1.

Table 1. Percentage comparison for Sample 1

AI Image Generation Tool	Similarity (%)
Gemini	60
Open AI (Dall E3)	70

II. Input Prompt for Sample 2: Draw for me reading glasses standing on a black table, black background glasses temple orange shield sand colour and round bridge blue front colour blue.




Original Image	Dall-E 3 Output Image	Gemini Output Image
 <p data-bbox="421 723 456 757">(a)</p>	 <p data-bbox="815 723 850 757">(b)</p>	 <p data-bbox="1222 723 1257 757">(c)</p>

Image2. Comparison of Image 1. (a) from both AI generation tools

Input Prompt #2 The first image is the real object. The second image is the image generated by Open AI(Dall-E 3). The third image is the image generated by Gemini. Now tell the differences between them specifically and tell their similarities in percentage. Which one is similar to the real object and what percentage?

Result and Analysis: Gemini generated appears more similar to the real object in terms of being eyeglasses rather than sunglasses and having a two-tone color scheme that includes orange.

Differences between Gemini and Dall- E Image Outputs: The first image displays eyeglasses with clear lenses, featuring a reflective blue-green tint on one lens and a dark-coloured frame with a subtle tortoiseshell pattern, positioned at an angle. The second image showcases sunglasses with dark lenses, a two-toned frame with a dark blue top and orange bottom, and a perfectly symmetrical front view, giving off a more stylized and polished appearance. Lastly, the third image presents eyeglasses with clear lenses, a navy blue frame with orange temples (side arms), and a perfectly symmetrical front view, exhibiting a stylized, illustrated appearance with gold outlines.

Similarities of Gemini and Dall E Image Output: Both Outputs were compared in percentage to analyze the similarity with the original image using [Claude-AI](#) (Claude, 2024) The images present individual variances, yet all showcase eyewear with circular frames. While each depiction is unique, the fundamental shape and concept remain consistent, demonstrating diverse interpretations. Similarity percentage is given in Table 2.

Table 2. Percentage comparison for Sample 2

AI Image Generation Tool	Similarity (%)
Gemini	70
Open AI (Dall E3)	60

III. Input Prompt for Sample 3: To me, it's a round acoustic guitar with an orange frame and a black exterior against a white wall.

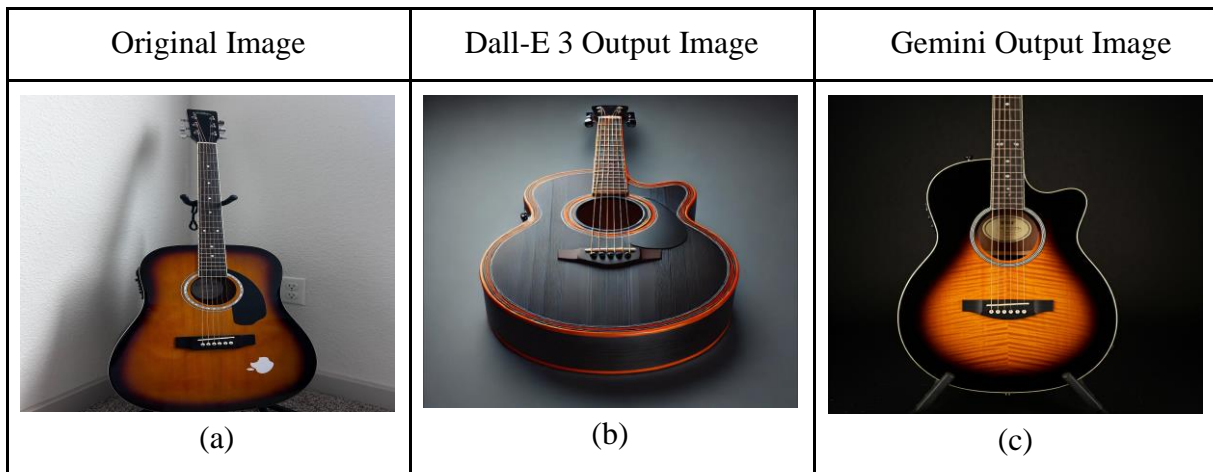


Image3. Comparison of Image 1. (a) from both AI generation tools

Input Prompt #3 The first image is the real object. The second image is the image generated by Open AI(Dall-E 3). The third image is the image generated by Gemini. Now tell the differences between them specifically and tell their similarities in percentage. Which one is similar to the real object and what percentage?

Result and Analysis: The Gemini image has a higher similarity rate compared to the Open AI-generated image. The similarity is mainly in the overall style and color harmony. However, the image generated by both (Gemini and Open AI) contains significant differences in terms of details and complexity compared to the better design of the real guitar.

Differences between Gemini and Dall- E Image Outputs: Image 2 has a black body with orange edges, while Image 1 has a sunburst finish. Image 2 shows a more elaborate, artistic design with curved edges and a non-standard shape. The sound hole in Image 2 is slightly off-centre and oval-shaped, as opposed to the circular hole in Image 1. Image 2 does not have the visible electrical components (magnets) seen in Image 1. Image 3 has a black top with sunburst edges, while Image 1 has a full sunburst finish. The body shape in Image 3 is slightly



different, with a more pronounced cutout. Image 3 shows the flame maple top, which is not in Image 1. The fret inlays in Image 3 are different.

Similarities of Gemini and Dall E Image Output: Both Outputs were compared in percentage to analyse the similarity with the original image using [Claude-AI](#) (Claude, 2024) All three images depict steel-string acoustic guitars. They all have similar general structures: body, neck, headstock, bridge, and soundhole. The overall proportions and placement of the components are largely similar. Similarity percentage is given in Table 3.

Table 3. Percentage comparison for Sample 3

AI Image Generation Tool	Similarity (%)
Gemini	85
Open AI (Dall E3)	70

IV. Input Prompt for Sample 9: Draw for me a grey Tesla toy car on a black background looking sideways over the edge.

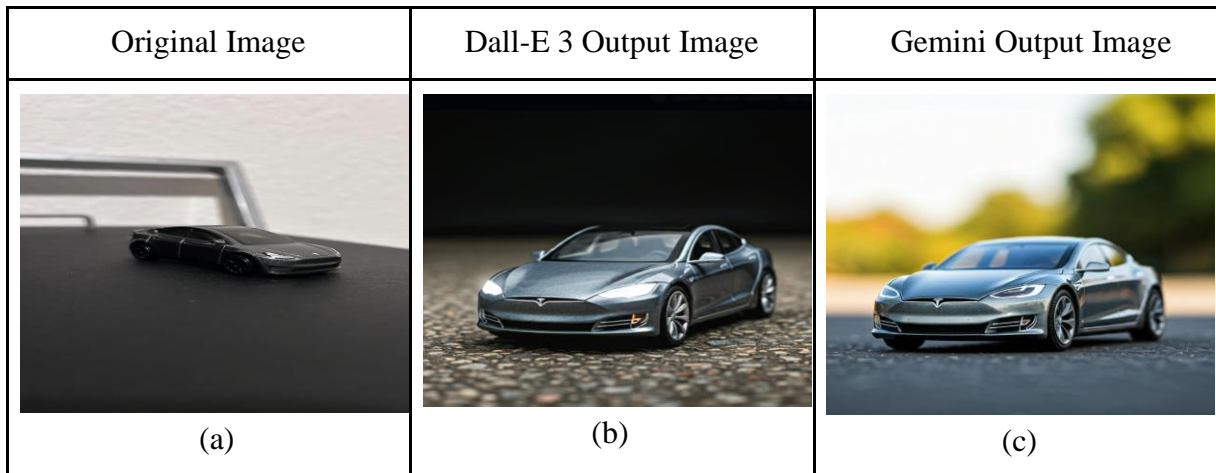


Image4. Comparison of Image 1. (a) from both AI generation tools

Input Prompt #4 The first image is the real object. The second image is the image generated by Open AI(Dall-E 3). The third image is the image generated by Gemini. Now tell the differences between them specifically and tell their similarities in percentage. Which one is similar to the real object and what percentage?

Result and Analysis: Which one is most similar to the real object: Image 2 (OpenAI generated) is more similar to the real object (Image 1) in terms of overall composition and lighting. However, it's important to note that Image 2 represents a much more detailed and larger scale model than the toy car in Image 1.

Differences between Gemini and Dall- E Image Outputs: Small, black toy car model. Appears on a dark surface against a light background. Less detailed, simpler model. Lighting is more even and less dramatic. A larger, more detailed Tesla Model S Metallic grey color model. Photographed in a low-light environment with dramatic lighting. More realistic proportions and details. Also a larger, more detailed Tesla Model S Metallic grey colour model. Photographed outdoors with natural light and a blurred background. Very realistic appearance with sharp details

Similarities of Gemini and Dall E Image Output: Both Outputs were compared in percentage to analyze the similarity with the original image using [Claude-AI](#) (Claude, 2024). All three images depict Tesla Model S cars, though at different scales and levels of detail. They all show the distinctive Tesla front fascia design and overall body shape. Similarity percentage is given in Table 4.

Table 4. Percentage comparison for Sample 4

AI Image Generation Tool	Similarity (%)
Gemini	55
Open AI (Dall E3)	60

V. Input Prompt for Sample 5: For me, it is a drawing of a black mouse with Logitech branding, drawn from above on a black table .

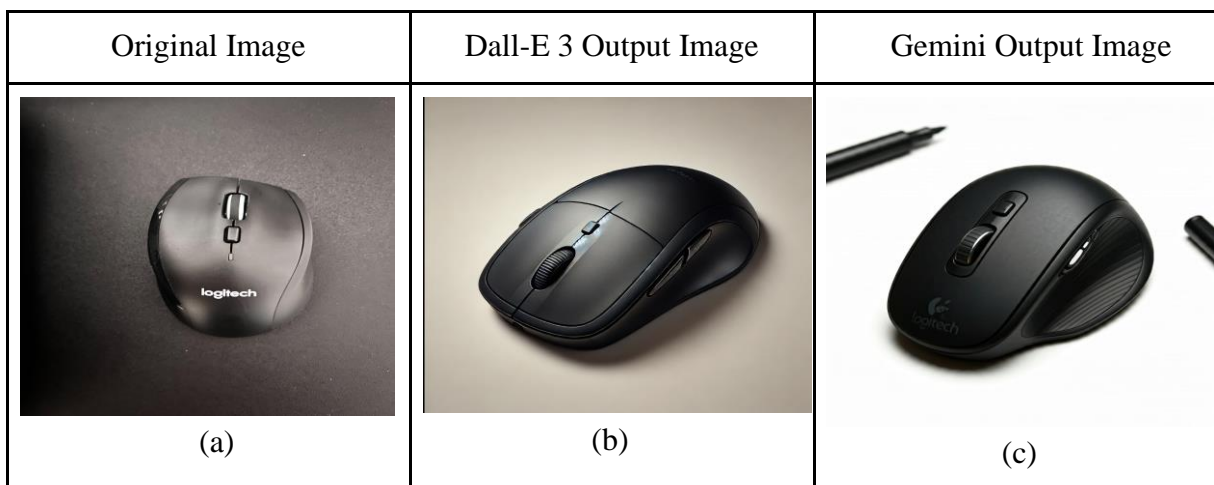


Image5. Comparison of Image 1. (a) from both AI generation tools

Input Prompt #5 The first image is the real object. The second image is the image generated by Open AI(Dall-E 3). The third image is the image generated by Gemini. Now tell the differences between them specifically and tell their similarities in percentage. Which one is similar to the real object and what percentage?

Result and Analysis: Shape and design:Image 1 (real): More compact and rounded shape .Image 2 (OpenAI): Slightly larger and more elongated , Image 3 (Gemini): Larger with a more pronounced ergonomic shape , Image 1: Two visible buttons (scroll wheel and side button) , Image 2: Three visible buttons (scroll wheel, left/right click, side button).

Differences between Gemini and Dall- E Image Outputs:Image3:Two visible buttons (scroll wheel and side button), but differently positionedBranding:Image 1: "logitech" logo clearly visible , Image 2: No visible branding , Image 3: "logitech" logo visible, but less prominent .Image 1: Smooth, matte finish . Image 2: Smooth, slightly glossy finishImage 3: Matte finish with some textured areas , Image 1: Dark, textured surface ,Image 2: Light, smooth surface Image 3: White background with a pen visible .

Similarities of Gemini and Dall E Image Output: scales and levels of detail. All three images depict a black computer mouse. The overall concept of a wireless mouse is consistent across all images . Similarity percentage is given in Table 5.

Table 5. Percentage comparison for Sample 5

AI Image Generation Tool	Similarity (%)
Gemini	55
Open AI (Dall E3)	60

VI. Input Prompt for Sample 6: Draw for me a white square Airpods Pro 5 model with a dark background on the black table.

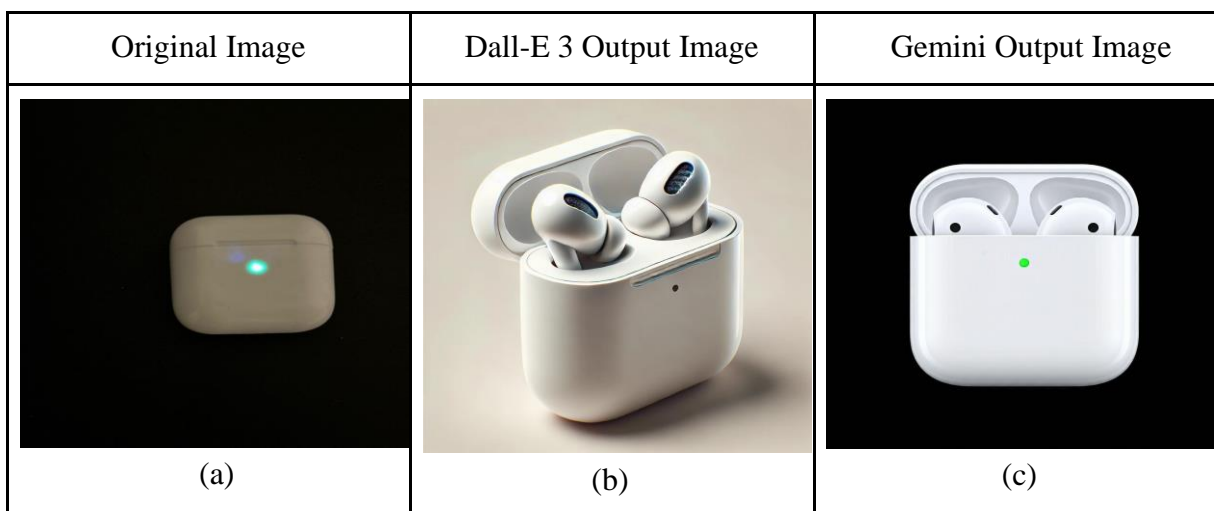


Image6. Comparison of Image 1. (a) from both AI generation tools

Input Prompt #6 The first image is the real object. The second image is the image generated by Open AI(Dall-E 3). The third image is the image generated by Gemini. Now tell the differences between them specifically and tell their similarities in percentage. Which one is similar to the real object and what percentage?

Result and Analysis: The Open AI image has a higher similarity rate compared to the image produced by Gemini. The similarity is mainly due to the overall style and colour matching. However, the image produced by both (Gemini and Open AI) contains significant differences in detail and complexity compared to the better design of the actual headset.

Differences between Gemini and Dall- E Image Outputs:Image 1 (real object) and Image 2 (generated by OpenAI):Image 2 shows the AirPods case open with the earbuds visible, while Image 1 shows only the closed case.Image 2 has more detailed shading and a 3D look, while Image 1 is a flatter case with less defined edges.The LED indicator in Image 2 is not visible, while Image 1 shows a green LED. Image 1 and Image 2: 70% similar :Image 1 and Image 3: 85% similar : Image 2 and Image 3: 75% similar .

Similarities of Gemini and Dall E Image Output: Overall, all three images depict an AirPods charging case, maintaining the general shape and colour scheme. They all capture the essence of the product.Similarity percentages (approximate):The most similar image is Image 3 (generated by Gemini), with an estimated 85% similarity. Similarity percentage is given in Table 6.

Table 6. Percentage comparison for Sample 6

AI Image Generation Tool	Similarity (%)
Gemini	55
Open AI (Dall E3)	60

VII. Input Prompt for Sample 7: Draw for me a dark iPhone 14 plus model phone dark background.




Original Image	Dall-E 3 Output Image	Gemini Output Image
 <p data-bbox="421 913 456 949">(a)</p>	 <p data-bbox="815 913 850 949">(b)</p>	 <p data-bbox="1220 913 1256 949">(c)</p>

Image7. Comparison of Image 1. (a) from both AI generation tools

Input Prompt #7 The first image is the real object. The second image is the image generated by Open AI(Dall-E 3). The third image is the image generated by Gemini. Now tell the differences between them specifically and tell their similarities in percentage. Which one is similar to the real object and what percentage?

Result and Analysis: The OpenAI-generated image (Image 2) is more similar to the real object (Image 1) in terms of overall shape and proportions. I estimate it to be around 80-85% similar to the real object.

Differences between Gemini and Dall- E Image Outputs: Image 1 (real object): Shows a matte black iPhone with dual cameras, Camera module is square-shaped, Apple logo is silver/white , Background is slightly textured . Image 2 (OpenAI generated): Shows a glossy black iPhone with a curved design ,Display is visible and illuminated with abstract shapes , No visible camera module , Angled perspective view .Image 3 (Gemini generated): Shows a glossy dark green iPhone with triple cameras ,Camera module is square with rounded corners, Apple logo is silver/white ,Background is completely black Similarities to the real object (Image 1)

Similarities of Gemini and Dall E Image Output: Similarities to the real object (Image 1):Image 2 (OpenAI): ~30% similar .Similarity percentage is given in Table 7.

Table 7. Percentage comparison for Sample 7

AI Image Generation Tool	Similarity (%)
Gemini	55
Open AI (Dall E3)	60

VIII. Input Prompt for Sample 8: Draw for me a black drone with an overflow model and a black background.

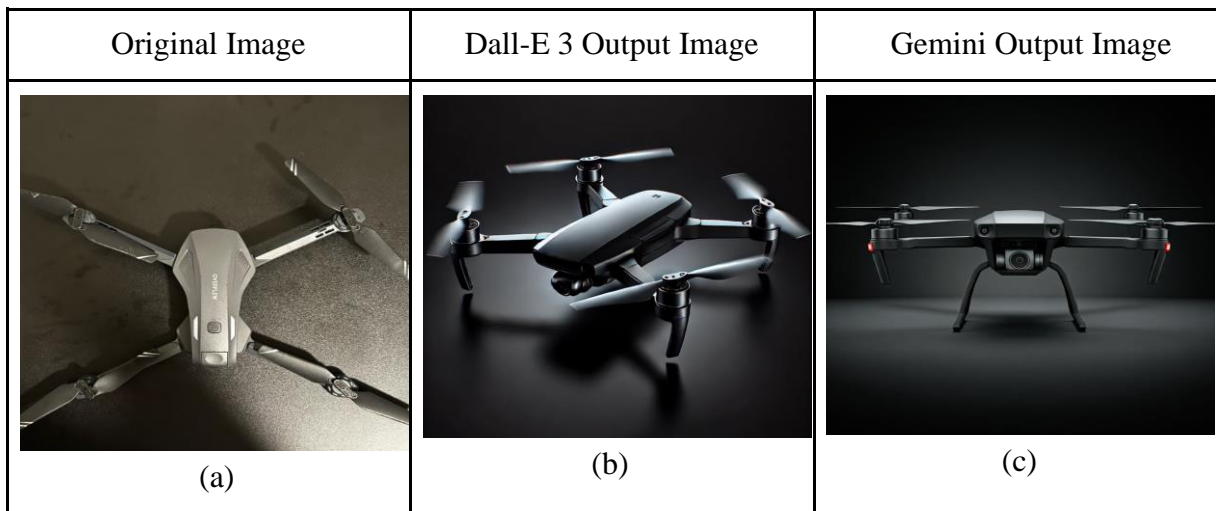


Image8. Comparison of Image 8. (a) from both AI generation tools

Input Prompt #8 The first image is the real object. The second image is the image generated by Open AI(Dall-E 3). The third image is the image generated by Gemini. Now tell the differences between them specifically and tell their similarities in percentage. Which one is similar to the real object and what percentage?

Result and Analysis: The OpenAI generated image (Image 2) is more similar to the real object (Image 1) in terms of overall shape and proportions. I estimate it to be around 80-85% similar to the real object.

Differences between Gemini and Dall- E Image Outputs: Image 1 (Real object): Shows a drone branded "PERFECT QQ" , Has a more basic, utilitarian design , Appears to be photographed on a textured surface , Visible screws and connections , Propeller blades clearly visible . Image 2 (rendered by OpenAI): Sleeker, more fluid design ,No visible branding ,Photographed on a reflective surface ,More professional studio-like lighting . Image 3 (rendered by Gemini): Most advanced design of the three , Features red LED lights ,Has a prominent camera module ,Photographed on a dark, gradient background Landing gear visible



Similarities of Gemini and Dall E Image Output: All three images depict quadcopter drones with similar general structures. They all have four propellers, a central body, and an aerodynamic design. The basic functionality and purpose of the devices are consistent across all images. Similarity percentage is given in Table 8.

Table 8. Percentage comparison for Sample 8

AI Image Generation Tool	Similarity (%)
Gemini	55
Open AI (Dall E3)	60

XI. Input Prompt for Sample 9: Draw for me a glass from the glass on the table.

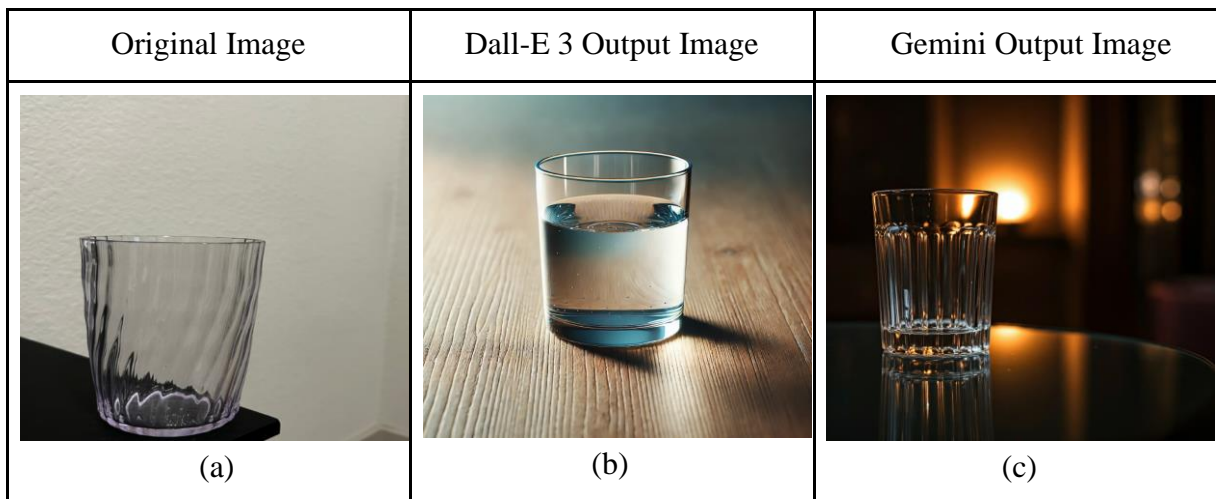


Image 9. Comparison of Image 1. (a) from both AI generation tools

Input Prompt #9 The first image is the real object. The second image is the image generated by Open AI(Dall-E 3). The third image is the image generated by Gemini. Now tell the differences between them specifically and tell their similarities in percentage. Which one is similar to the real object and what percentage?

Result and Analysis: Shape and design: Image 1 (real): The glass has a swirled, diagonal pattern and a slightly wider base. Image 2 (OpenAI): The glass is perfectly cylindrical with no pattern. Image 3 (Gemini): The glass has vertical ridges/flutes.Color: Image 1: The glass has a slight purple tint.

Differences between Gemini and Dall- E Image Outputs: Image 1: The glass has a slight purple tint. Image 2: The glass is completely clear. Image 3: The glass is clear but reflects warm light. Content : Image 1: Empty , Image 2: Filled with water , Image 3: Empty ,Background and lighting: Image 1: Plain light background, shadows visible . Image 2: Wooden table, soft lighting . Image 3: Dark background with warm light source

Similarities of Gemini and Dall E Image Output: Image 1 vs Image 2: Approximately 70% similar Image 1 vs Image 3: Approximately 80% similar . Similarity percentage is given in Table 9.

Table 9. Percentage comparison for Sample 9

AI Image Generation Tool	Similarity (%)
Gemini	55
Open AI (Dall E3)	60

5. Conclusions

This study highlights the strengths and weaknesses of the main AI fashions, Gemini and OpenAI (DALL-E 3), in generating high-best snap shots primarily based on activate enter. Both fashions display the potential to create visually distinct photographs. But there may be a noticeable distinction in coloration, accuracy, detail, and object recognition. Excellent shade fidelity and satisfactory detail decision This makes it best for creating particularly accurate and sensible photographs. OpenAI's DALL-E three, on the other hand, stands proud for its ability to interpret complex prompts.

Overall, the pictures generated by using OpenAI were observed to be more similar to the unique gadgets. For Gemini, you'll get 70% versus 60% suits; however, both fashions have their own strengths. Depending on the specific use case, Gemini is a good choice for tasks that require high individual precision and excellent detail. At the same time, OpenAI is able to perform better in situations that require concise and accurate descriptions.

Ultimately, this vision provides creators and users with valuable insights into selecting the right model for their visualisation responsibilities. Because those AI devices are constantly evolving.

References

- Blitzstein, J. K., & Hwang, J. (2015). *Introduction to probability*. Crc Press.
- Bartram, S. M., Jürgen Branke, & Mehrshad Motahari. (2020). *Artificial Intelligence in Asset Management*. CFA Institute Research Foundation.
- Claude. (2024). Claude.ai. <https://claude.ai/login?returnTo=%2F%3F>
- Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press. (Original work published 2021)
- D. Jason Slone, & McCorkle, W. W. (2019). *The Cognitive Science of Religion*. Bloomsbury Publishing.
- Virvou, M., Tsihrintzis, G. A., Tsoukalas, L. H., Jain, L. C., & Springerlink (Online Service. (2022). *Advances in Artificial Intelligence-based Technologies : Selected Papers in Honour of Professor Nikolaos G. Bourbakis--Vol. 1*. Springer International Publishing, Imprint Springer.
- Ford, M. (2018). *Architects of intelligence : the truth about AI from the people building it*. Packt Publishing.

