The Analysis of "Dall-E 2" Digital Image Outcome on "Starry Night, Van Gogh" Prompt.

Bayu Setiawan[1], Pungky Febi Arifianto[2], Damara Alif Pradipta[3]
[1,2]UPN Veteran Jatim
[3]UISI
E-mail: [1]bayusetiawan.dkv@upnjatim.ac.id,
[2]pungkyarifianto.dkv@upnjatim.ac.id,
[3]damara.pradipta@uisi.ac.id

Abstract

The evolution of AI (artificial intelligence) is massive and limitless nowadays. Dall-E 2 was launched limitedly on 20 July 2022. This is an interesting subject to pay attention to since the Dall-E 2 Developer claims that the program will give more realistic and aesthetic results than the older version. Dall-E 2 is a digital image generator software developed by Open AI. This program works by converting (generating) text called "prompt" into a digital image. Digital images generated by Dall-E2 claimed that have unlimited possibilities, where the results will be very different, and no single image will be the same. In this study. We did a limited analysis of the generated outcome imagery-specific prompts. This research is about visual analysis of "Starry Night" text prompts generated images outcome. The experimentation on 28 prompts indicates that the generated images had fine quality in image detail and accuracy. The image outcome is diverse in color, consistency, composition, object, and style. Dall-E 2 image outcome is great even it has limitation and flaw.

Keywords: AI generator, Dall-E 2, art, generative art, AI

INTRODUCTION

The evolution of AI (artificial intelligence) is massive and progresive nowadays. There is more and more AI technology that is introduced nowadays. In the scope of visual art and design, AI works by generating pictures and video from simple sources. There is several AI that has been taking people's attention, just like Disco Diffusion, Starry AI, Wombo ART, Midjourney, and DALL-E.

Dall-E is one of the AI programs that have been popular since it released their latest version called DALL-E 2. Dall-E 2 was launched limitedly on 20 July 2022. This is an interesting subject to pay attention to since the Dall-E 2 Developer claims that the program will give more realistic and aesthetic results than the older version. Dall-E 2 is a digital image generator software developed by Open AI.

"kittens playing with yarn in a sunbeam"

Image 1. The image outcome from DALL-E2
Source: Swimmer, 2022

Dall-E 2 is really easy to use since it works on computer browser. This program works by converting (generate) text called "prompt" into digital images. The developer of the program claims that the digital images generated by Dall-E 2, have unlimited possibilities, where the results will be very different, and no single image will be the same.

Image 2. Dall-E 2 image outcome from prompt
"An astronaut riding a horse in a photorealistic style"
Source: https://openai.com/dall-e-2/

Ordinary people can easily create professional-looking artworks using AI programs, which results in more diversity in the creative expressions of society. Lee (2022) Said that artworks generated by AI may not be aesthetically impressive, but they could be much more affordable and accessible than works by professional

artists. What Lee said can be irrelevant today based on the latest AI technology. So here we to test Dall-E 2 image outcome to know how good AI works today.

MATERIALS AND METHOD

To do this research, we are using Gillan Rose Visual methodologies mage analysis theories. This research focuses most strongly on the image itself, and although it pays the most attention to this compositionality, it also pays some attention to its production (Rose: 2001).

In this research, we only can use 28 prompts due to prompt limitations. In case focusing on the prompt's usage, we are limiting the prompt to the main subject and the art style.

Each Prompt in Dall E 2 produces 4 images outcome. As a result, we have 112 images of 28 to use as material. We are grouping ce each prompt to analyze each image by the visual aspect. To analyze the image, we set codes about based on Gillian Rose visual methodologies. We set 3 codes about content, colour, and spatial organization, The prompts that wes use ia about "Borobudur" and "starry Night" and sort of artstyle.

In this test, we test DALL-E2 to combine the main object "borobudur temple" with "Starry Night" Van Gogh painting and additional style. Wes adds several styles just lika cartoon, ghibli, art nouveau, art deco, psycadelic, manga, disney, cyberpunk, etc.

FINDINGS/RESULTS

Huang (2022) in his research found that DALL-E works by combining the most distinctive and recognizable features of the subject and the style based on their image's library. These "features" may include facial characteristics, costumes, hairstyles, makeup, accessories, color palettes, brushstrokes, modeling of light and shadow, compositions, lines and shapes, etc. The next question is how big their library is, and how they are combining several specific subjects and style in one prompt.

Image 3. The real images of Borobudur temple (left)and Starry Night painting (right)
Source: borobudur (Left), https://www.indonesia.travel, Starry Night (Right),
https://www.vangoghgallery.com

For the first test DALL- E 2 give a great result to generate simple prompt about Borobudur and Starry Night. When we add more spesific style, DALL-E 2 show a great result by object representation. After we tried more style, we found that DALL-E 2 had a different style library database. The result of the prompt looks similar after we try several styles. All of the image outcome looks great, detail and eyecatching.
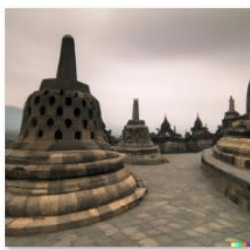
content	The object represented accurately, no deformation, realistic photo
colour	Natural colour based on real object
spatial organization	Pointed to the *stupa.* Eye level

DALL·E 2022-09-27 20.22.37 - borobudur temple realistic painting.png

DALL·E 2022-09-27 20.22.40 - borobudur temple realistic painting.png

DALL·E 2022-09-27 20.22.42 - borobudur temple realistic painting.png

DALL·E 2022-09-27 20.22.45 - borobudur temple realistic painting.png

content	The object represented accurately, no deformation, water colour
colour	Natural colour based on real object
spatial organization	Pointed to the top *stupa* alligntment. Low Eye level

DALL·E 2022-09-27 20.08.29 - starry night van gogh.png

DALL·E 2022-09-27 20.08.41 - starry night van gogh.png

DALL·E 2022-09-27 20.08.49 - starry night van gogh.png

DALL·E 2022-09-27 20.08.57 - starry night van gogh.png

content	The sky represented great detail, the land various landscape, deformation cloud and star, oil paint colour,
colour	Dominate by black and dark blue, Contrast star color yellow
spatial organization	Focus on the sky and stars. Normal Eye level

DALL·E 2022-09-27 20.09.51 - starry night van gogh realistic painting.png

DALL·E 2022-09-27 20.09.55 - starry night van gogh realistic painting.png

DALL·E 2022-09-27 20.09.58 - starry night van gogh realistic painting.png

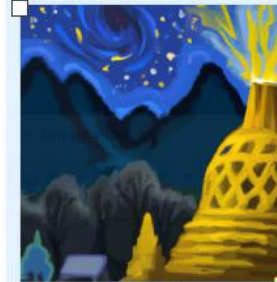DALL·E 2022-09-27 20.10.02 - starry night van gogh realistic painting.png

content	The sky represented great detail, the land various landscape, oil paint colour
colour	Dominate by black and dark blue, Contrast star color yellow
spatial organization	Focus on the sky and stars. High Eye level

DALL·E 2022-08-08 12.23.37 - a cartoon van gogh starry night borobudur realistic painting.png

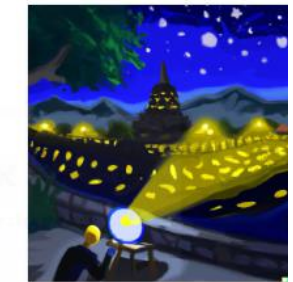DALL·E 2022-08-08 12.23.43 - a cartoon van gogh starry night borobudur realistic painting.png

DALL·E 2022-08-08 12.23.47 - a cartoon van gogh starry night borobudur realistic painting (1).png

DALL·E 2022-08-08 12.23.47 - a cartoon van gogh starry night borobudur realistic painting.png

content	The sky represented great detail, Deformed borobudur, Additional object added, oil paint colour
colour	Dominate by black and dark blue, Contrast yellow star color, Yellow Lamp, and temple
spatial organization	Focus on the temple, lamp, and star. Normal Eye level

| DALL·E 2022-08-08 13.06.52 - a high resolution photo van gogh starry night borobudur realistic painting.png | DALL·E 2022-08-08 13.06.58 - a high resolution photo van gogh starry night borobudur realistic painting.png | DALL·E 2022-08-08 13.07.07 - a high resolution photo van gogh starry night borobudur realistic painting.png | DALL·E 2022-08-08 13.07.11 - a high resolution photo van gogh starry night borobudur realistic painting.png |

content	The sky represented low detail, Deformed borobudur, additional tree object, oil paint colour
colour	Dominate by black and dark blue, Contrast yellow star color, Yellow Lamp, and temple
spatial organization	Focus on the sky, temple, and star. Normal Eye level

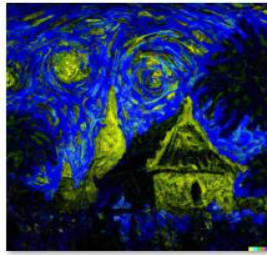| DALL·E 2022-09-28 07.47.57 - a van gogh starry night borobudur realistic photo.png | DALL·E 2022-09-28 07.48.00 - a van gogh starry night borobudur realistic photo.png | DALL·E 2022-09-28 07.48.02 - a van gogh starry night borobudur realistic photo.png | DALL·E 2022-09-28 07.48.05 - a van gogh starry night borobudur realistic photo.png |

content	The sky represented great detail, Deformed borobudur, Additional object added, oil paint colour
colour	Dominate by black and dark blue, Contrast yellow star color, Dark Grey temple
spatial organization	Focus on the temple, and star. Normal Eye level, High eye level

DALL·E 2022-08-08 13.07.55 - a psychedelic van gogh starry night borobudur realistic painting.png

DALL·E 2022-08-08 13.08.01 - a psychedelic van gogh starry night borobudur realistic painting.png

DALL·E 2022-08-08 13.08.05 - a psychedelic van gogh starry night borobudur realistic painting.png

DALL·E 2022-08-08 13.08.08 - a psychedelic van gogh starry night borobudur realistic painting.png

content	The sky represented great detail, Deformed borobudur, Additional object added, oil paint colour
colour	Dominate by black yellow and dark blue, Contrast yellow star color, Yellow Lamp house and temple
spatial organization	Focus on the temple, lamp and star. Normal Eye level, High eye level

DISCUSSION

Recognizable Uniqueness

In His Research, Huang (2022) found that DALL-E 2 generates art by combining the most distinctive and recognizable features of the subject and the style. These "features" may include facial characteristics, costumes, hairstyles, makeup, accessories, color palettes, brushstrokes, modeling of light and shadow. DALL-E 2 Maybe had a different image library for object and style. In makes DALL-E 2 cannot generate unique images for several prompts. In some images DALL-E 2 looks cannot generate objects perfectly that makse the objects look weird and illogical. Even each image of is different and unique, DALL-E 2 we can recognize the images is not "man made" by several flaw in the image that make it looks so "artificial".

Boundaries of Unlimited possibilities

Limited unlimited possibilities are something we found in DALL-E 2. There are boundaries that make DALL-E 2 can not freely generate the images. Even in Huang (2022) reseach DALL-E 2 show has different levels of art style comprehension. It has a greater understanding of postmodern and contemporary art styles, especially digital art and some cartoon styles linked to popular animations. That result is little bit different in this research. For certain prompt, DALL-E 2 cannot generate it right.

DALL-E 2 cannot generate the style of "art deco', art neuvou". The image outcome looks similar and has no specific artstyle like the prompt.

Chaotic representation

Mazzone (2019) in his research found that the art generated will be novel but at the same time will not depart too much from acceptable aesthetic standards. This is called the "least effort" principle in Martindale's theory, and it is essential in the art generation because too much novelty would result in rejection by viewers. In this research "least effort" also makes the images looks so weird and illogical. First is the representation of borobudur, there are almost 1 of 2 images of Borobudur is deformed into different building. Some of the additional objects also make it look bad, some of the subjects added in certain promt. But it does not look good and makes it more looks "Chaotic".

CONCLUSION

The development of AI is getting faster and smarter. As a designer and visual artist, it can be a treat and opportunity. Dall-E has been doing a great jump in AI development in these 4 years. For the first program use, the imager outcome looks great and cool, but after several trials, the result is becoming monotone. Dall E 2 is giving a great result of the image generator program today. Even though it has several flaws, Dall E 2 produces a great image outcome

REFERENCES

Rose, Gillian. (2007). *Visual Methodologies: An Introduction to the Interpretation of Visual Materials*. SAGE.

Ettinger, A. (2020). What BERT is not: Lessons from a new suite of psycholinguistic diagnostics for language models. *Transactions of the Association for Computational Linguistics*, 34-48.

Marcus, G. F., & Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust*. Pantheon

Monro, G. (2009). Emergence and Generative Art. *Leonardo*, 42(5), 476–477. doi: 10.1162/leon.2009.42.5.476Press.

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., . . . Sutskever, I. (2021). Learning transferable visual models from natural language supervision. [Proceeding]. International Conference on Machine Learning, (pp. 8748-8763).

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). *Hierarchical Text-Conditional Image Generation with CLIP Latents*. arXiv preprint arXiv:2204.06125. Retrieved from https://arxiv.org/abs/2204.06125

Rips, L. (1989). *Similarity, typicality, and categorization. In S. Vosniadou, & A. Ortony, Similarity and Analogical Reasoning (pp. 21-59).* Cambridge: Cambridge University Press.

Singh, Gautam. Fei Deng & Sungjin Ahn. (2022). Illiterate Dall-E Learns To Compose. Rutgers University. KAIST. [Proceeding]. Published as a conference paper at ICLR 2022

Swimmer963. (2022, May 1). *What DALL-E 2 can and cannot do. Retrieved from LessWrong*: https://www.lesswrong.com/posts/uKp6tBFStnsvrot5t/what-dall-e-2-can-and-cannot-do

Thrush, T., Jiang, R., Bartolo, M., Singh, A., Williams, A., Kiela, D., & Ross, C. (2022). *Winoground: Probing Vision and Language Models for Visio-Linguistic Compositionality*. arXiv preprint arXiv:2204.03162. Retrieved from https://arxiv.org/abs/2204.03162

Kim, Jina, Soyeon Shin, Kunwoo Bae, Soyoung Oh, Eunil Park, and Angel P. del Pobil. (2020). "Can AI Be a Content Generator? Effects of Content Generators and Information Delivery Methods on the Psychology of Content Consumers." *Telematics and Informatics 55* (December 2020): 101452. https://doi.org/10.1016/j.tele.2020.101452.

Lee, Hye-Kyung. (2022). "Rethinking Creativity: Creative Industries, AI and Everyday Creativity." *Media, Culture & Society 44,* no. 3 (March 7, 2022): 601–12. https://doi.org/10.1177/01634437221077009.

Mazzone, Marian, and Ahmed Elgammal. (2019). "Art, Creativity, and the Potential of Artificial Intelligence." *Arts* 8, no. 1 (February 21, 2019): 26. https://doi.org/10.3390/arts8010026.

Monro, Gordon. "Emergence and Generative Art." *Leonardo* 42, no. 5 (October 2009): 476–77. https://doi.org/10.1162/leon.2009.42.5.476.

Web

https://www.indonesia.travel/content
https://www.vangoghgallery.com/painting/starry-night.html)
https://openai.com/dall-e-2/