

A Research on Aesthetic Aspects of Checkpoint Models in [Stable Diffusion]

Ke Ma, Jeanhun Chung*

*Doctor's Course, Dept. of Multimedia, Graduate School of Digital Image and Contents
Dongguk University, Korea*

**Professor, Dept. of Multimedia, Graduate School of Digital Image and Contents
Dongguk University, Korea*

*a1044894679@gmail.com, * evengates@gmail.com*

Abstract

The Stable diffusion AI tool is popular among designers because of its flexible and powerful image generation capabilities. However, due to the diversity of its AI models, it needs to spend a lot of time testing different AI models in the face of different design plans, so choosing a suitable general AI model has become a big problem at present. In this paper, by comparing the AI images generated by two different Stable diffusion models, the advantages and disadvantages of each model are analyzed from the aspects of the matching degree of the AI image and the prompt, the color composition and light composition of the image, and the general AI model that the generated AI image has an aesthetic sense is analyzed, and the designer does not need to take cumbersome steps. A satisfactory AI image can be obtained. The results show that Playground V2.5 model can be used as a general AI model, which has both aesthetic and design sense in various style design requirements. As a result, content designers can focus more on creative content development, and expect more groundbreaking technologies to merge generative AI with content design.

Keywords: AI Generated Technology, AI Painting, Generative Artificial Intelligence, 3D Content Design, Stable Diffusion, Checkpoint Models

1. Introduction

1.1 Research Background

Stable Diffusion as the hottest generative AI painting technology, due to its open source characteristics and can be deployed locally in the computer qualities, loved by the majority of designers. Stable Diffusion in the SD1.5 or SDXL base model as a baseline, to use their own different styles of pictures to train the model, but

Manuscript Received: April. 26, 2024 / Revised: May. 4, 2024 / Accepted: May. 11, 2024

Corresponding Author: evengates@gmail.com (Jeanhun Chung)

Tel: +82-2-2260-3, Fax: +82-2-2260-3766

Professor, Dept. of Multimedia, Graduate School of Digital Image and Contents, Dongguk University, Korea

Copyright© 2024 by The Institute of Internet, Broadcasting and Communication. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>)

due to the different use of everyone's designs, the results of the trained model vary, and the different models of the cue word understanding varies.

However, due to the different design purposes of each person, the results of the trained models are not the same, and the understanding of the cue words varies from model to model. Whenever designing AI images in a certain style, we need to switch between different models to meet the design requirements. In this paper, we introduce the use of two different Stable Diffusion models to generate different types of image styles, and by comparing the AI images generated by the two models, we analyse their aesthetics from the perspectives of the match between the image and the prompt, and the surface composition and colour composition of the image, to select a general AI model that meets the various design requirements, and the AI images generated by this model have a certain aesthetic appeal.

1.2 Research Purpose

The aim of the study is to test the use of two of the most popular Stable Diffusion generative AI models, and to summarise the aesthetic strengths and problems in the images by analysing the characteristics of the AI-generated images. The aim of this study is to explore and analyse AI image generation models that meet various design needs by investigating two AI image generation models for content designers to design image content.

2. Theoretical Background

2.1 Stable Diffusion Model

Checkpoint is one of the most important models in Stable Diffusion, and is the master model on which almost all operations depend. Checkpoint models are trained on Stable Diffusion models, and are therefore sometimes referred to as Stable Diffusion models.

The Checkpoint model plays a key role in Stable Diffusion. It is a model that has been trained and optimised to generate and infer content such as text, images, etc. The Checkpoint model can be used for a variety of tasks such as text generation, image repair, image generation, etc. The Checkpoint model can be used for a variety of tasks such as text generation, image repair, and image generation. By using Checkpoint models, users can generate a wide variety of content according to their needs.

The training of Checkpoint models is done by using large scale datasets. During the training process, the model learns how to understand and generate text, images, and other content. Through iterative training, the model optimises its capabilities and improves the quality and diversity of the content it generates.

3. Stable Diffusion Model AI Image Aesthetics Comparison

In this study, two Stable Diffusion models were used, the most popular recent Stable Diffusion model "Playground V2.5" selected from the Civitai website and the Stable Diffusion Juggernaut XL model. In ComfyUI, we tested the aesthetic differences between the AI images generated by these two models using a basic Venn diagram workflow. We chose a girl as the main object of comparison, and paired it with different styles of images to compare the performance, strengths and weaknesses of the two AI models in terms of the planar and colour composition of the generated images.

3.1 Character Style Comparison

First of all the base prompt with four different styles to generate the relevant AI images, choose five different kinds of objects as the base prompt, no negative respectively using the Juggernaut XL model and the Playground V2.5 model, the CFG of the Juggernaut XL model is set to 2.0, Steps The CFG of Juggernaut XL model is set to 2.0, Steps is set to 6, and the sample of dpmpp_2m_sde is used; the CFG of Playground V2.5 model is set to 3.0, Steps is set to 40, and the sample of euler is used. and 4 different AI images are generated at one time, which ensures that the difference in the image quality due to the error in the parameters is minimized.

As shown in Figure 1, the AI images of the characters are compared, from the generated AI images, both Juggernaut XL model and Playground V2.5 model can understand the input prompt well, and generate the corresponding AI images strictly according to the prompt. The character images generated by the Juggernaut XL model tend to be photorealistic, while the character images generated by Playground V2.5 model have been processed by specialized art. Juggernaut XL model generates character images that tend to be more photorealistic, while Playground V2.5 model generates character images that have been processed with specialized artwork, resulting in better light and shadow performance, and a better overall aesthetic sense of the image.



Figure 1. Characters Image Comparison

As shown in Figure 2, the comparison of AI images of buildings, from the generated AI images, Juggernaut XL model and Playground V2.5 model can understand the input prompt very well, and generated the corresponding AI images strictly according to the prompt. The color of the images generated by the Juggernaut XL model is more monotonous, and does not have any sense of aesthetics. The images generated by the Juggernaut XL model are monotonous in color and do not have any aesthetics. The images generated by the Juggernaut XL model are not clearly structured, and the elements in the images are too monotonous, whereas the images generated by the Playground V2.5 model not only have a better understanding of the prompts, but also have a better composition and better color performance.

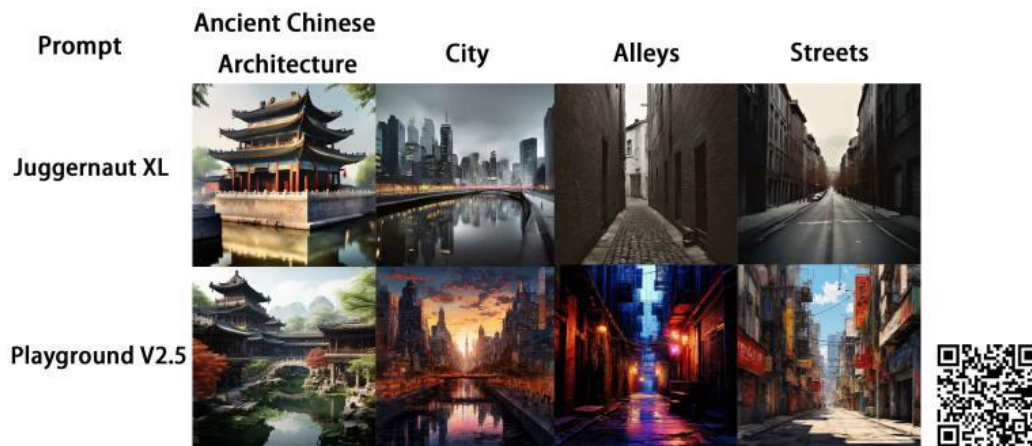


Figure 2. Buildings Image Comparison

As shown in Figure 3, the comparison of AI images of food, from the generated AI images, both Juggernaut XL model and Playground V2.5 model can understand the input prompt well, and generate the corresponding AI images strictly according to the prompt, the only difference is that in the image representation of "Barbecue", Juggernaut XL generates the image of barbecue, probably the shape is not a big problem, if you look closely, the shape of barbecue is blurred, and it is not possible to identify the type of object. The only difference is that in the image representation of "Barbecue", the image of barbecue generated by Juggernaut XL has little problem with the approximate shape of the barbecue, but when looking closely, the shape of the barbecue is blurred and the type of the object is not recognisable.



Figure 3. Food Image Comparison

As shown in Fig. 4, the comparison of AI images of animals, from the generated AI images, both Juggernaut XL model and Playground V2.5 model can understand the input prompt well, and generated the corresponding AI images strictly according to the prompt. Both AI models can generate common animals in reality as well as fantasy creatures very well. However, Playground V2.5 is better than Juggernaut XL in terms of composition and lighting, and the overall color contrast is stronger and more aesthetically pleasing.



Figure 4. Animals Image Comparison

.As shown in Figure 5, the AI images of the carrier are compared, from the generated AI images, the Playground V2.5 model can understand the input prompt well, and generates the corresponding AI images strictly according to the prompt. The image generated by the Juggernaut XL model is only an approximate shape match for the prompt, and when we look closely, it doesn't match the prompt and there are cases of interspersing between objects. On closer inspection, the images do not match the prompt, and there are interspersed objects. On the other hand, the vehicle images generated by the Playground V2.5 model not only match the shape and structure of the Prompt, but also have a better color composition.



Figure 5. Carrier Image Comparison

4. Conclusion

The comparative analysis of the above cases concludes that the Playground V2.5 model outperforms the AI images generated by the Juggernaut XL model in terms of overall style and detail, and generates images that are more aesthetically pleasing and design-oriented, with stronger colour contrasts in the colour compositions, and more closely related styles in terms of the details of the generated images. This study explores AI models that can be generalised to any design scenario by comparing the images generated by the two AI models as the basis of the study. Based on the results of this study, it appears that the Playground V2.5 model can be used as a general AI model for both image aesthetics and prompt matching in stable diffusion, a generative AI tool that can better help content designers to develop more interesting and creative content.

References

- [1] Ke Ma, Chung, Jeanhun Chung. "A Research on 3D Texture Production Using Artificial Intelligence Softwear " The International Journal of Advanced Culture Technology Vol.15 No.4 178-184,2023
DOI:<https://doi.org/10.7236/IJIBC.2023.15.4.178>
- [2] Biying Guo, Xinyi Shan, Jeanhun Chung. "A Comparative Study on the Features and Applications of AI Tools - Focus on PIKA Labs and RUNWAY " The International Journal of Internet, Broadcasting and Communication Vol.16 No.1 86-91,2024 DOI:<https://doi.org/10.7236/IJIBC.2024.16.1.86>
- [3] Chen Xi, Jeanhun Chung. "A Study on Character Design Using [Midjourney] Application" The International Journal of Advanced Culture Technology Vol.11 No.2 409-414,2023
DOI:<https://doi.org/10.17703/IJACT.2023.11.2.409>
- [4] Yifan Cui, Xinyi Shan, Jeanhun Chung. "A Feasibility Study on RUNWAY GEN-2 for Generating Realistic Style Images" The International Journal of Internet, Broadcasting and Communication Vol.16 No.1 99-105,2024
DOI:<https://doi.org/10.7236/IJIBC.2024.16.1.99>
- [5] Chenghao Wang, Jeanhun Chung. "Research on AI Painting Generation Technology Based on the [Stable Diffusion] " The International Journal of Advanced Smart Convergence Vol.12 No.2 90-95,2023
DOI:<https://doi.org/10.7236/IJASC.2023.12.2.90>
- [6] Qianqian Jiang, Jeanhun Chung. "A Case Study of Creative Art Based on AI Generation Technology" The International Journal of Advanced Smart Convergence Vol.12 No.2 84-89,2023
DOI:<http://dx.doi.org/10.7236/IJASC.2023.12.2.84>
- [7] Pingjian Jie, Jeanhun Chung, Xinyi Shan. "A Comparative Analysis Between <Leonardo.Ai> and <Meshy> as AI Texture Generation Tools" The International Journal of Advanced Culture Technology Vol.11 No.4 333-339,2023
DOI:<https://doi.org/10.17703/IJACT.2023.11.4.333>
- [8] Pingjian Jie, Xinyi Shan, Jeanhun Chung. "Comparative Analysis of AI Painting Using [Midjourney] and [Stable Diffusion] - A Case Study on Character Drawing" The International Journal of Advanced Culture Technology Vol.11 No.2 403-408,2023 DOI:<https://doi.org/10.17703/IJACT.2023.11.2.403>
- [9] Ke Ma, Chung, Jeanhun Chung. "A Research on 3D Texture Production Using Artificial Intelligence Softwear " The International Journal of Advanced Culture Technology Vol.15 No.4 178-184,2023
DOI:<https://doi.org/10.7236/IJIBC.2023.15.4.178>
- [10] Stable Diffusion Art, How does Stable Diffusion work? https://stable-diffusion-art.com/how-stable-diffusion-work/#Diffusion_model