Fuzzy Systems and Data Mining IX
A.J. Tallón-Ballesteros and R. Beltrán-Barba (Eds.)

doi:10.3233/FAIA231089

Innovative Approach to Children's Desk Lamp Using Stable Diffusion Technology

Tianxin Li[1]
School of the Art Institute of Chicago, *USA*

Abstract. The emergence of AI intelligent drawing technology has completely revolutionized the traditional model of the design industry. With the help of smart technology, designers can quickly generate a large number of creative images in various styles and content. This greatly enhances design efficiency and has a revolutionary significance in promoting design innovation. This article focuses on the creative process in the design of children's table lamps. It uses the AI drawing software "Stable Diffusion Web UI" based on deep generation technology as a tool to combine the professional skills of designers with AI drawing technology to achieve the creative design goal of personalized children's table lamps. The specific content and methods include: (1) building a lamp image library, quantitatively describing the design style of children's table lamps, extracting composition, and summarizing the different structural-semantic features of lamps. (2) Generating feasible creative lamp product ideas based on the Stable Diffusion Web UI software. Through research and exploration, the style description and structural characteristics of children's table lamps were summarized, and the parameters for generating creative lamp ideas using the Stable Diffusion Web UI software were determined. This has been helpful in furthering the integration of artificial intelligence and product design.

Keywords. Children's desk lamp, creative approach, deep generation technology, stable diffusion.

1. Introduction

The creative process is a critical component of product design. For a long time, the design of the creative process heavily relied on the divergent thinking and manual drawing abilities of designers. However, traditional methods of manually creating creative images are limited by cognitive constraints, as well as the personal drawing styles and skills of the designers. This makes it difficult to achieve stylistic variations and innovations. At the same time, the inefficiency of manual creative drawing cannot meet the demands of rapid design iteration, creating bottlenecks in the development of the design industry.

The emergence of AI intelligent drawing technology has completely revolutionized the traditional model of the design industry. It significantly enhances design efficiency, which is revolutionary in promoting design innovation. However, the question remains: how can intelligent technology be harnessed by designers to better serve design goals? What are the different applications of intelligent drawing technology in various scenarios? Designers often encounter challenges during the application process and need to continually address and resolve them.

[1] Corresponding Author, Tianxin Li, School of the Art Institute of Chicago, USA; Email: litianxin2014@gmail.com

2. Related works

In recent years, the application of Artificial Intelligence (AI) technology across various domains has garnered significant attention. Researchers such as Gao Xuchao [1] and Li qinggang[2] have delved into the application of AI technology in railway construction and operation, while Zhao Tianqi[3]and others have explored the profound impact AI technology will have on the engineering field, particularly in fine-tuned design contributions. Wang Xiaohui[4] and his team have focused on AI's application in personalized cultural design, with an emphasis on the use of deep generation models.

Cheng Sun[5] and colleagues have applied AI technology in architectural design, discussing its practical application in the field from three perspectives: information integration, mapping modeling, and decision support. This provides a new perspective on how to use AI effectively in design. Sun[6] and his team have introduced a human-machine collaborative drawing system based on Generative Adversarial Networks (GAN) in their research, which can be used to generate highly realistic images. Further research by Ma Yongjie[7] and others categorizes and analyzes the improvements in GAN, showcasing its innovations and accomplishments in the field of AI drawing. Lianglei[8], Zhan Shi[9], and their peers explore the potential of SD models in generating general images. Zheng Kai and Wang Di's[10] research examines the application and development of AI image generation from a modular perspective, providing comprehensive insights into the relevant usage scenarios, methods, and instructions.

Dehouche Nassim[11] and others evaluate the potential of AI drawing in teaching art history, aesthetics, and techniques from an art education perspective. The British company Arom has successfully designed and produced candle products using their AromAI, signifying the vast application potential of AI technology in product design[12].

In summary, in the currently published papers, although researchers have made some progress in the area of AI-generated design, their research results and quantity are not sufficient, and most of the research areas remain in the field of architecture and engineering. The research on artificial intelligence technology in the field of product design is still weak.

This paper focuses on exploring the application of SD technology in creative product sketching. Taking specific goals and applications as a starting point, the study compares the problems, causes, and solutions of different SD technology approaches in realizing specific task goals. The study is a tentative exploration of SD technology in the field of product design. The research process and conclusions provide some references for the application of AI applications in the creative design stage of products.

3. Relevant theories and methodologies

3.1 Denoising Diffusion Probabilistic Models (DDPM)

DDPM is the foundation of the Stable Diffusion model, tasked with producing images that accurately reflect the training data's distribution. Its process involves feeding a text description and a random image into the model, which, upon learning, yields an image that aligns with the provided text. Figure 1 provides a visual depiction.

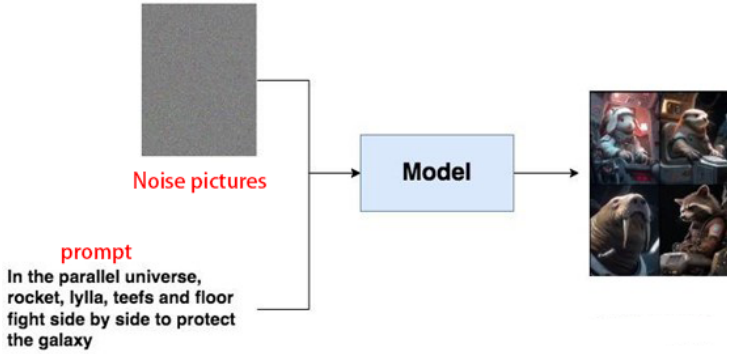

Figure 1 DDPM Image Generation Principle

3.2 Stable diffusion model

The Stable Diffusion Web UI is a drawing software based on the Stable Diffusion (SD) model. The SD model, also known as the Stable Diffusion model, is an advanced version of the DDPM model. When the SD model is used to create or edit an image, it first maps the image to a lower-dimensional space called the latent space. Compared to the original image space, this space has fewer dimensions, which greatly improves computational efficiency. The core of the SD model is the diffusion process, which includes both forward and backward diffusion.

In the forward diffusion process, the model gradually adds noise to the original image, causing the original features of the image to become progressively blurred until it finally appears as an image resembling random noise. The main focus of the diffusion model is on the backward diffusion process. In reverse diffusion, starting from a noise image without any features, the model gradually removes the noise and guides the generation process to eventually recover the original image.

To achieve this goal, the model must compute an "optimal noise" at each step, which indicates the amount of noise to be removed in the current step. This calculation can be effectively performed using a deep learning model, such as the noise predictor UNet. During the training process, the noise predictor continuously adjusts its weights to improve prediction accuracy. In the reverse diffusion process, using the predicted "optimal noise", the noise is gradually removed and the original image is recovered from the noisy image. Through repeated execution, a clear original image can be obtained.

In summary, the SD model utilizes both forward and backward diffusion processes as well as the predictive ability of the noise predictor to effectively recover the original image from a noisy image. In addition, the SD model library allows users to select and customize their training, resulting in the creation of model libraries with specific styles, such as realistic models. This high level of control has opened up new possibilities in the field of AI art and image generation.

3.3 Introduction to Sampling Methods

During the image generation process, users need to select different sampling methods to remove excess noise when generating images. The Stable Diffusion web interface integrates several samplers, most of them created by Katherine Crowson and based on research by Jiaming Song, Karras, Cheng Lu, and others. This article focuses primarily on the following sampling methods

(1) DPM2 by Karras

The DPM2 method uses ancestral sampling, and the setting of the "eta" parameter can affect the results.

(2) DPM++ 2S by Karras

This sampling method is based on the research of Cheng Lu and others. It uses K-diffusion technology to achieve second-order one-step operations and combines ancestral sampling. The parameter "eta" also affects the results in the DPM++ 2S a sampling method. Cheng Lu has made the implemented code available on GitHub, and users can customize their choices for first-order, second-order, third-order, and single-step or multi-step operations as needed. The web interface uses a version with default parameters.

(3) DPM++ 2M

This is a second-order multi-step sampling method based on the research of Cheng Lu and others, implemented under K-diffusion technology.

(4) LMS of Karras

This method, based on Karras' research, applies Karras' noise schedule method and shows strong advantages in terms of image color and vibrancy.

3.4 ControlNet

ControlNet is a neural network architecture that can control the SD model, allowing the SD model to accommodate a wider range of input conditions. The original SD model could accept input from prompts and source images, while ControlNet provides various input conditions, including canny edges, semantic segmentation maps, key points, and doodles, thus expanding the boundaries of SD's capabilities and greatly enhancing the controllability of AI art creation. The visual structure of ControlNet is shown in Figure 2.

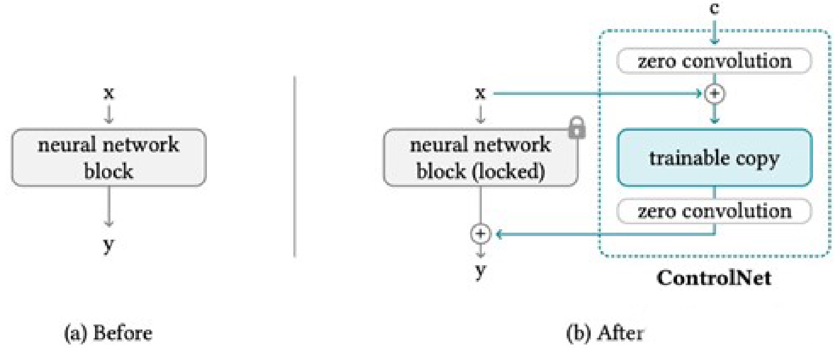

Figure 2 ControlNet's visual structure diagram.

The network structure of ControlNet can be divided into two parts: the "locked copy" and the "trainable copy". The "locked copy" fixes the original weights of the stable diffusion, preserving the image generation capabilities already learned by the stable diffusion.

In Figure 2, the zero convolution is a 1x1 convolutional layer with weight and bias initialized to zero. Since the zero convolution initializes both weight and bias to zero before training begins, the entire ControlNet model remains in its original stable diffusion state. Only after training with your own data will the trainable copy in Figure 2(b) and the learnable parameter values in the zero convolution change, allowing the entire network to learn the specific task you specify. In short, users can influence and control the generated images by uploading their own data, such as product structure images.

4. Experimental Exploration

This section provides a detailed description of the experimental process, the problems and challenges encountered, and how these problems were solved.

4.1 Experimenting with the Standard Stable Diffusion Model

In the initial phase of the research, this thesis chose the default stable diffusion model SD 1.5.safetensors as the starting point for the experiments. The experiments were conducted on a Windows 10 operating system with WebUI version ID 1.60. The goal of the experiment was for the model to generate images that meet the requirements of the experimenter, based on given keyword prompts.

Since the target audience is children, who tend to prefer cute product styles, keywords such as "desk lamp," "animals," and "children" were chosen as input in this paper. Several sampling methods were chosen for the sampling process, including DPM++ 2S a Karras, DPM2 a Karras, LMS Karras, and DPM++ 2M Karras. A step size of 30 was chosen, the CFG scale was set to 8, and the model was used to generate images corresponding to the given scenes (see Figures 3-6).

Figure 3 DPM2 a Karras Sample Generation

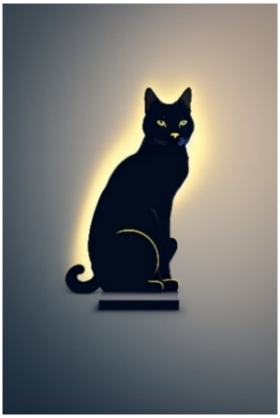

Figure 4 DPM++ 2S a Karras Sample Generation

Figure5 DPM++ 2M Karras Sample Generation

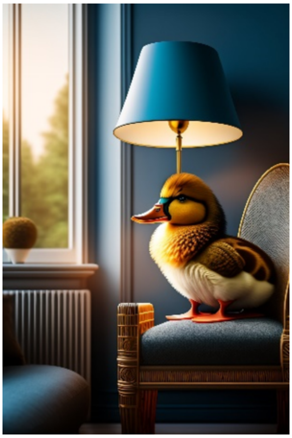

Figure 6 LMS Karras Sample Generation

From the above text, it can be seen that the images generated by the SD 1.5.safetensors model have a certain artistic quality in terms of visual style, but their performance does not meet expectations. For example, the generated image in Figure 3 has a strong artistic quality, but does not correspond to normal human perception. Figure 4 lacks the key element of a desk lamp, and shows only light and an animal. Figure 5 lacks a sense of design and does not meet the author's expectations for the animal-shaped desk lamp, showing only separate elements of an animal and a desk lamp. Figure 6 lacks a logical connection, and the connection between the animal and the desk lamp seems rigid and cannot be used as a creative image.

The research results show that SD 1.5.safetensors can accommodate different styles and are suitable for designers trying to generate images for the first time. However, its disadvantage lies in the lack of control over the style and logical coherence of the generated images, which requires operators to make detailed adjustments.

Regarding the sampling method, based on experiments, LMS Karras produces the best results in generating creative product images among the four sampling methods. Despite a strong artistic quality, the generated products have bright colors, relatively good texture, and logical coherence compared to the other sampling methods. Therefore, LMS Karras is chosen as the sampling method for the following experiments.

In terms of descriptions, the performance of the standard SD 1.5.safetensors model is quite poor. In most cases, it cannot correctly interpret the intended meaning of the text expressed by the author. This is especially obvious when it comes to generating images that combine two nouns, e.g. when the author enters "animal lamp", as the generated image often shows separate entities of an animal and a lamp. According to the analysis, the reason for this is that the standard training dataset of the stable diffusion model consists mainly of images with an artistic style, which results in images that often show pronounced fusion features. In addition, the model clearly lacks the necessary training for more complex product image generation tasks.

4.2 Experimental approach using the realistic model

Given the limitations of the default stable diffusion model, SD 1.5.safetensors, this experiment aims to explore alternative solutions. Specifically, we're attempting creative image generation using the advanced Realistic model, which is based on SD

1.5.safetensors. In addition, we will optimize and modify the input prompts, including reverse prompts, to improve the image generation process. Some examples of reverse prompt keywords are "blur", "low resolution", "text", "crop", "worst quality", "low quality", "normal quality", "jpeg", "artifacts", "signature", "watermark", "username", "blur", "text", "signature", "watermark", "simple background", "cartoon", "date", "low resolution", "line art", "flat color", and so on.

For the sampling method, we use LMS Karras. In addition, we've chosen the "Latent (nearest)" option in the upscaling algorithm to perform high-resolution enhancement of the images. This process involves 10 iterations, a redraw magnitude of 0.7, and a magnification factor of 2. These settings are used for image generation. In this experiment, we also use different product style keywords such as "Chinese style," "antique style," "simple style," and "European style" to guide and diversify the image generation trials (see Figures 7-10).

Figure 7 Chinese style

Figure 8 antique style

Figure 9 simple style

Figure 10 European style

The research results show that by using the realistic model and adjusting the parameters while optimizing the prompts, the generated images show improved logical coherence and visual quality compared to the standard model. However, there are still limitations in terms of structure; the model struggles to produce fine and precise results. For example, there are logical inconsistencies in the position of the light bulb design, structural anomalies in the animals (such as the head appearing in the chest area and strange proportions), and illogical relationships between the desk lamp and the animal (where the lamp is positioned over the animal and the animal is treated as a decorative element).

An analysis of the generated images suggests that the realistic model relies more on random processes for generation, indicating its shortcomings in detail control. As a result, it can be concluded that the Realistic model, compared to the standard Stable Diffusion model, is more inclined to a realistic style in its training dataset. Therefore, it is more suitable for generating everyday objects.

4.3 ControlNet Introduction Experiment

To address the problems of the unreasonable animal structure in the creative image generation of the above product, the irrational relationship between the animal and the table lamp, and the illogical positioning of the bulb design, the ControlNet plugin was introduced in this experiment. This plugin provides more maneuverability for AI drawing, for example, it can clearly show the image structure generated by the model through a simple schematic, as shown in Figure 10. The author used the ControlNet plugin to draw the desired product structure features - the lower body is elliptical, the upper head is smaller relative to the lower part, and their relationship is superimposed. At the same time, the keywords were set as an animal table lamp, children, minimalist style, and the lamp inside the animal. ControlNet used the perfect pixel mode, with the control type set to segmentation, the preprocessor set to reference_only, the control weight set to 1, the guided access set to 0.28, the termination timing set to 1, and the control mode set to balanced. Image generation was then performed. After filtering, the creative image of the product that met the author's expectations was obtained (see Figure 11).

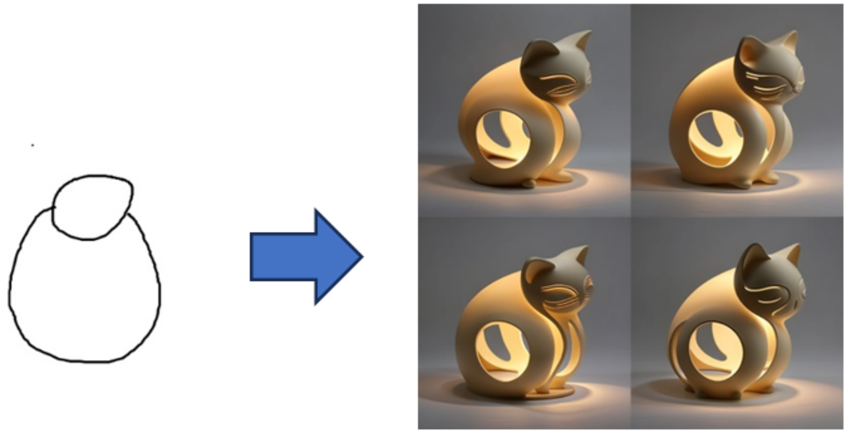

Figure 11 The partial product creative image generated after adjustments through ControlNet

The research has shown that by using the ControlNet plugin, the generated images can be more precisely controlled in terms of content and structure by the author, allowing for more accurate image generation. This is a significant step forward. The ControlNet plugin not only addresses previous issues with model-generated structures but also provides a more flexible and controllable approach to generation. It has proven to be a great help in product design and creative image generation. The author predicts that with further optimization of the ControlNet plugin's options and prompts, the AI will be able to better understand human drawing intentions and thus generate more accurate images that meet specific requirements.

5. Conclusion

The research of this paper is based on AI drawing technology, exploring its potential and limitations in the field of product design and creative image creation. Through experimental research using the Stable Diffusion original model, the Realistic model, and the ControlNet plugin, it explores the possibilities of AI-generated product creative images while discussing their practicality and limitations.

Through the experimental research, it was found that in generating product creative images, designers can use the Realistic model as a basic model. At the same time, they should use simple descriptive vocabulary as positive prompts and terms such as blur, low resolution, text, cropping, worst quality, low quality, normal quality, JPEG artifacts, signatures, and watermarks as negative prompts. Furthermore, when the prompt is clearly defined along with the desired product structure, using the ControlNet plug-in to generate creative images yielded the best results.

The experimental research also revealed the importance of selecting and fine-tuning the appropriate AI models and providing precise inputs (such as the ControlNet plug-in's schematic and fitting parameters) when using AI drawing to generate product creative images.

This study validates to some extent the potential application of AI technology in the field of product design and creative image creation. It also reveals some challenges and problems that may be encountered in practical operation, such as logical problems with the internal structure of products and machine misinterpretation of prompts. Subsequent research will further explore and develop the application of AI in product creative image generation, focusing on improving the accuracy of creative image generation and the feasibility of generating product internal structures.

Reference

[1] Gao Xuchao. Research on Engineering Brain Technology Based on Artificial Intelligence and Federated Learning. Railway Construction Technology, 2023(3), 1-4, 13.
[2] Li Qinggang. Research on the Technology Framework and Development Direction of Intelligent Metro. Railway Construction Technology, 2022(11), 52-56.
[3] Zhao Tianqi, Gou Hongye, Chen Xuanying, et al. Research Progress on Bridge Informatization and Intelligent Bridges in 2020. Journal of Civil and Environmental Engineering (in Chinese and English), 2021, 43(S1), 268-279.
[4] Wang Xiaohui, Qin Jingyan, and Quan Hongchen. "Personalized Cultural Creative Product Design Method Based on AI Artwork Generation." Packaging Engineering, 41.06 (2020): 7-12. doi:10.19554/j.cnki.1001-3563.2020.06.002.
[5] Sun Cheng, Han Yunsong, and Ren Hui. "Research on Architectural Computational Design for Artificial Intelligence." Architectural Journal, 2018(9), 98-104.
[6] SUN Lingyun, CHEN Pei, XIANG Wei, et al. "SmartPaint: A Human-Machine Collaborative Drawing System Based on Generative Adversarial Neural Networks." Frontiers of Information Technology & Electronic Engineering, 2019, 20(12), 1644-1657.
[7] Ma Yongjie, Xu Xiaodong, Zhang Ru, et al. "Research Progress on Generative Adversarial Networks and Their Application in Image Generation." Computer Science and Exploration, 2021, 15(10), 1795-1811.
[8] Liang L. Exploration and Improvement of the Stable Diffusion Model in the Field of Image Generation. Advances in Computer and Communication, 2023, 4(3).
[9] Shi Z. AI Application to Generate an Expected Picture Using Keywords with Stable Diffusion. Journal of Artificial Intelligence Practice, 2023, 6(1).
[10] Zheng Kai, Wang Di. Application of Artificial Intelligence in the Field of Image Generation - Using Stable Diffusion and ERNIE-ViLG as Examples. Technology Vision, 2022, 35, 50-54.

[11] Nassim D, Kullathida D. What's in a text-to-image prompt? The potential of stable diffusion in visual arts education. Heliyon, 2023, 9(6).
[12] AromAI Launches First Candle Designed Entirely by AI: The Future of Fragrance is Here. M2 Presswire, 2023.