GEN AI AND INTERIOR DESIGN REPRESENTATION: APPLYING DESIGN STYLES USING FINE-TUNED MODELS

Hyun Jeong, Youngchae Kim & Youngjin Yoo
Dept of Interior Architecture and Built Environment, Yonsei University, South Korea

Seunghyun Cha
Korea Advanced Institute of Science and Technology (KAIST), South Korea

Jin-Kook Lee
Dept of Interior Architecture and Built Environment, Yonsei University, South Korea

ABSTRACT: *This paper explores the applicability of Image-generation AI in the field of interior architectural design, with a particular focus on automating interior design representation based on design styles. Interior design representation involves a complex process that integrates visual elements with functionality and user experience. Effectively visualizing this process is essential for facilitating communication among the various stakeholders involved in the design process. However, traditional visualization methods are constrained by expert resources, costs, and time limitations. In contrast, image-generation AI has the potential to automate various design elements, including design styles, components, and spatial arrangements, to enhance representation. In this study, we evaluated the performance of a base model using various design styles and, based on the evaluation results, selected styles for fine-tuning. The methodology for fine-tuning these design styles involved the following steps: 1) data preparation and preprocessing, 2) hyperparameter optimization, and 3) model training and construction. Utilizing the fine-tuned model thus constructed, we conducted image generation demonstrations. The research results revealed that design styles not well represented by the base model were effectively captured, and high-quality images were generated by the fine-tuned model. Notably, this fine-tuned model demonstrated the ability to represent images of specific design styles with a high degree of accuracy in capturing the characteristics and keywords associated with each style, compared to the base model. This implies that through fine-tuning image-generation AI, a wide range of applications can be inferred when aiming to create customized designs by considering these aspects. In conclusion, this study explores an efficient approach to interior design representation in the field of interior architecture by employing image-generation AI and proposes a method to effectively generate visualized images by training on design style keywords. Through this approach, our study can contribute to improving the interior design process by facilitating the generation of visualized images that reflect design styles. Furthermore, the study aims to suggest the potential for applying this approach not only to the field of interior architecture but also across various domains to achieve effective visualization.*

KEYWORDS: *Interior Architecture Design, Interior Design Representation, Generative AI, Model Fine-tuning*

1. INTRODUCTION

Interior design representation plays a crucial role in the field of interior architecture, effectively conveying ideas and designs through visual media and facilitating effective communication among various stakeholders involved in the design process (Chiu, 1995). In interior spaces, design styles signify the approach and method of planning and decorating a space, shaping and emphasizing the aesthetic, functional, and psychological aspects of the space. Design styles encompass a variety of preferences and trends influenced by the users and purposes of the space, impacting choices in color, patterns, materials, furniture, and accessories (Goldschmidt et al., 1998; Eckert et al., 2000). Additionally, they serve as a means to reflect individual identity and lifestyle, reflecting personal preferences and tastes.

Therefore, understanding and proposing customized designs that consider user preferences in the spatial visualization process is essential. However, this process necessitates expertise to comprehend the diverse preferences and requirements of users, as well as the desired design styles and spatial elements for visualization. This requires a significant investment of time, cost, and effort for both experts and non-experts (Lee et al., 2020).

Recent advancements in deep learning technology have sparked significant interest in generative artificial intelligence (Gen AI). As a result, various research endeavors are underway in the realm of visual content creation using image-generation AI (Image-Gen AI) based on large language models (LLMs). Expanding upon this trend, our study aims to propose an approach for automating interior design representation using Image-Gen AI. This

Referee List (DOI: 10.36253/fup_referee_list)
FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)
Hyun Jeong, Youngchae Kim, Youngjin Yoo, Seunghyun Cha, Jin-kook Lee, *Gen AI and Interior Design Representation: Applying Design Styles Using Fine-Tuned Models,* pp. 950-957, © 2023 Author(s), CC BY NC 4.0, DOI 10.36253/979-12-215-0289-3.95

approach allows for the generation of diverse design visualization alternatives based on user preferences and objectives, all without the need for specialized expertise.

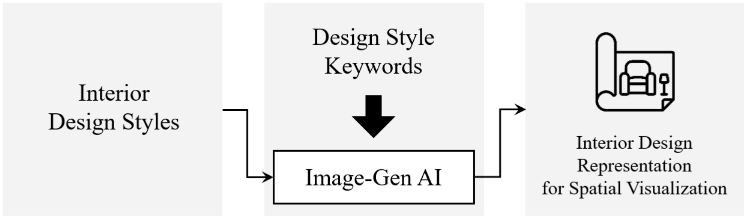


Fig 1: An Overview of the Study

2. BACKGROUND

2.1 Deep Learning-based Image-Gen AI

Image-Gen AI is based on deep learning and is a versatile technology applicable in various fields, including natural language understanding, computer vision, image processing, data generation, prediction, and more (Liu et al., 2021) This technology is used to generate new image content or outputs based on given data or information. To accomplish this, Image-Gen AI is pre-trained on large datasets and then fine-tuned for specific targets. During the training phase, Image-Gen AI learns features and patterns from the data, and in the generation phase, it uses this learned information to generate new image data. This process can be considered an example of transfer learning, allowing effective generation results even with limited data. Image-Gen AI can be trained to incorporate additional conditions such as image type, style, color, and more, enabling it to generate images that meet specific criteria. This reduces the need for extensive training on the target while enabling various applications like style transformation, ensuring similarity between images, image synthesis, and more (Nichol et al., 2021).

For these reasons, recent applied research efforts are being conducted in the field of visual content creation using a variety of Image-Gen AI models such as Midjourney (Oppenlaender, 2022), DALL-E 2 (Ramesh et al., 2022), Stable Diffusion, and others (Ramesh et al., 2022; Saharia et al., 2022; Rombach et al., 2022; Oppenlaender, 2022). While research using Image-Gen AI has been extensive, it remains limited in the field of interior architecture. Therefore, in this study, we aim to explore an approach to fine-tune Image-Gen AI models based on design styles and implement a model for the automatic visualization of interior design representation.

2.2 Potential for Automating Interior Design Representation through Image-Gen AI

In the field of interior architecture, the evolution of deep learning technology is reshaping the way spaces are conceptualized and realized. In the past, designers relied on manual sketches and 2D drawings to convey ideas for spatial visualization reflecting interior design representation (Ching, 2011). However, with the emergence of advanced technology and computer-aided tools, spatial visualization has undergone a paradigm shift (Karras et al., 2018). The integration of sophisticated software, computer-aided design, and 3D modeling tools has empowered designers to visualize spaces realistically and immersively. Thanks to these advancements, designers can accurately represent intricate details such as lighting, materials, textures, and shadows, not just the physical layout. As a result, stakeholders, including clients and project collaborators, can experience the proposed design in a lifelike manner before actual construction commences (Ah-soon & Tombre, 1997; Oxman, 2006).

With the continuous advancement of deep learning technology, Image-Gen AI can effectively generate images that match the intended target by fine-tuning a base model pretrained on large datasets. Leveraging these characteristics of Image-Gen AI, it is possible to implement an interior design representation model based on specific design keywords, fine-tuning it to reflect the desired design style. This model can be utilized as a tool to generate a variety of design alternatives that users desire during the design process and enhance the decision-making process (Jeong & Lee, 2023). Therefore, in this study, we aim to conduct fine-tuning of design styles on Image-Gen AI and explore methods for automating interior design representation.

3. MODEL FINE-TUNING FOR INTERIOR DESIGN REPRESENTATION

3.1 Overall Process

In this study, the image generation performance of three major AI platforms in the Image-Gen AI field, Stable Diffusion, DALL-E 2, and Midjourney, was examined and compared. First, these Image-Gen AI platforms employ two main methods: text-to-image (Txt2img) and image-to-image (Img2img). Each of these platforms has unique and distinctive image generation capabilities, along with their specific technical attributes, strengths, and limitations.

DALL-E 2, trained on a large-scale image dataset, demonstrates exceptional abilities in generating detailed and complex images. However, due to its complexity and resource-intensive nature, it may have longer processing times, and its dependency on text prompts might result in shortcomings in image stability and consistency. Midjourney excels in generating images inspired by specific visual styles or artistic aesthetics. However, its emphasis on artistic expression may result in relatively less accurate representation of real objects or scenes. Stable Diffusion prioritizes stability and accuracy in image generation. It is based on LLMs and provided as an open-source platform, making it user-friendly for customization. The transparency of the source code allows users to understand, modify, and apply the underlying algorithms, enabling them to verify, align with their intended purposes, and enhance image generation results.

Each platform has its unique strengths and specific limitations. Considering these factors, this study utilized Stable Diffusion, which demonstrated the most outstanding performance in terms of image generation stability, accuracy, and the ability to cater to specific requirements through open-source access. The research methodology, rooted in the utilization of Stable Diffusion, is structured into three primary phases. Step 1: Testing image generation via text descriptions for design styles, Step 2: Model Fine-tuning, Step 3: Evaluation of Fine-tuned Models. The schematic representation of the entire process outlined in this section is depicted in Fig 2 below.

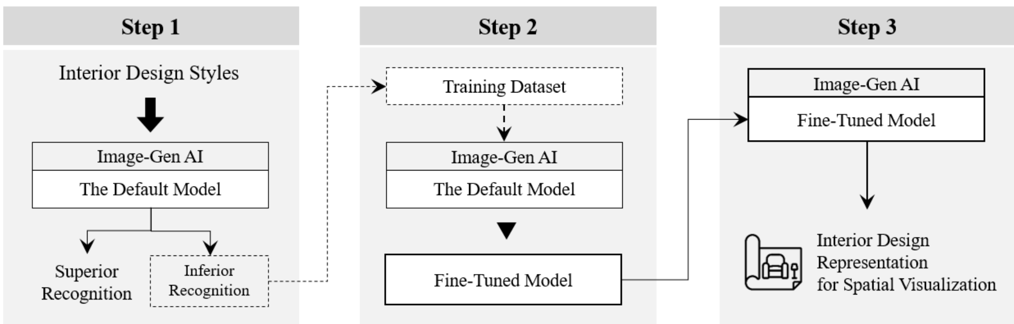


Fig 2: The Process of Model Fine-tuning

3.2 Step 1: Testing Image Generation via Text Descriptions for Design Styles

In this section, we conduct tests to evaluate the performance of the Image-Gen AI model for various interior design styles. Our study focuses on the text-to-image (txt2img) approach, generating images based on text descriptions. To prevent the Image-Gen AI model from inferring styles solely from text descriptions, we perform prompt engineering adhering to guidelines. This prompt consists of two key components: 1) prompts for the target design style, and 2) prompts for image quality. Additionally, we categorize positive and negative aspects that encompass both reflective and non-reflective elements of the generated images.

The image generation process utilized the DPM+2M Karras sampler along with the widely used open-source model, SD1.5V checkpoint (v1-5-pruned.ckpt). Essential configurations, including sampling steps and CFG scales, were set to default values, and the image size was defined as 1024x512 pixels. Each image generation took an average processing time of approximately 5 seconds. Table 1 below illustrates the configuration settings employed in the image generation process, and Table 2 showcases the standard format of the text prompts utilized for image generation.

Based on the preceding discussion of the image generation process, we conducted an evaluation of the recognition level for interior design styles using the generated images. Table 3 below provides outcomes of the generated images based on their recognition levels. The base model of Stable Diffusion demonstrated a stable generation capability for high-quality images across most styles. However, its expressive capacity was relatively limited in terms of being recognized as specific design styles, particularly due to lower comprehension of certain styles. To address these constraints and enhance image generation accuracy, we deduce that fine-tuning for specific targets is imperative.

Table 1: Configuration Settings for Image Generation

GPU	Base model	Sampling Method	Sampling steps	CFG Scale	Resolution
A6000 47.5 VRAM	SD v1.5 ckpt	DPM+ 2M Karras	25	13	1024×512 (2:1)

Table 2: The Standard Format of Prompts Used for Image Generation

Prompts	Positive	Negative
Design Style	Design Style interior, Space zoning	*None*
Image Quality	Professional photograph, photorealistic rendering, realistic, enhance-detail, v ray rendering, full HD, masterpiece, highly detailed, high quality, 8k, full shot, deep depth of field, f/22, 35mm	Bad proportion, Low quality, awkward shadows, unrealistic lighting, pixelated textures, Worst, noisy, unrealistic reflections, normal quality, watermark

Table 3: Sample Results of Generated Images based on Recognition Levels

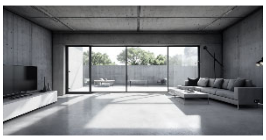
Recognition Level	Superior	Inferior
Image-Gen AI based generated Images	 "*Industrial* Style Interior, A living room"	 "*Brutalism* Style Interior, A living room"

3.3 Step 2: Model Fine-Tuning for Design Styles

In this section, the process of fine-tuning the model for the target design style (ex. Brutalism) encompasses three key steps: 1) Data Preparation, 2) Hyperparameter Optimization, and 3) Training.

Firstly, in the Data Preparation step, we focused on the collection and preprocessing of data tailored to the specific target design style. This Training Dataset requires two primary components: Image Data and Text Data. Table 4 provides an example of the training dataset.

Table 4: An example of the training dataset / ***content*** (*e.g.*, a space), ***style*** (*e.g.*, a design style) and *scene description*.

Image Data	Text Data
	"A ***Brutalism style*** interior in a ***Living room*** *with sharp lines that exemplify the brutalist aesthetics. Featuring shades of grey, concrete, and metallic tones, it showcases a minimalist living room characterized by grey hues and a monochrome color scheme."*

The subsequent stage involved the optimization of the model's hyperparameters to elevate its image generation performance for the specified design styles. This process encompassed the refinement of parameters like learning rate, batch size, and network architecture to attain improved outcomes. Table 5 presents the hyperparameters employed during the model fine-tuning.

Table 1: Optimized Hyperparameters for Model Fine-tuning

Training data	Epoch (Training steps)	Batch size to train	Learning rate	Learning rate Scheduler	Learning rate warmup
15	100	1	0.0001	Constant	10

The ultimate training phase involved the model being subjected to training using the meticulously prepared dataset and meticulously optimized hyperparameters. The process of model fine-tuning was executed using the training dataset and specified hyperparameter configurations to create the target model. The training duration amounted to approximately 25 minutes. Through this training, the model was tasked with comprehending the distinct attributes and intricacies of the designated design style, consequently empowering it to craft images that exhibit a heightened alignment with the intended aesthetic. This fine-tuned model is utilized in conjunction with the default model during subsequent image generation endeavors.

3.4 Step 3: Evaluation of Fine-tuned Models

Through the subsequent fine-tuning process, the model learned from both image and text data related to design styles that initially exhibited inferior recognition. As a result, it was enhanced to effectively depict the distinctive characteristics of the trained design styles, showcasing a heightened ability to generate high-quality images. Table 6 below illustrates the results of image generation based on the application or absence of the fine-tuned model. This highlights the tangible impact and comparison of interior design style image generation that was not achievable before the model's fine-tuning adjustments. Additionally, the weights of the model parameters correspond to finer adjustments made to the model, reflecting a more distinct influence on the image generation process.

Table 6: Qualitative Comparison of Image Generation: Impact of Applying Fine-Tuned Model

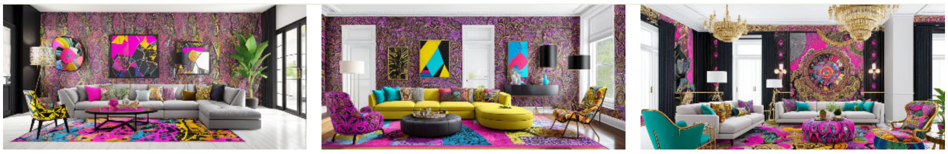
Weight of Fine-Tuned Model	at W 30%	at W 60%	at W 90%
Image-Gen AI based generated Images			
	"*Brutalism* Style Interior, A living room"		

4. DEMONSTRATION

In this study, based on the procedures outlined in the previous Section 3, a demonstration was conducted using Image-Gen AI to generate images of more than 15 interior design styles. The target space was limited to residential living rooms. The generated image results for each design style are presented in Table 7, encompassing both the default model and the fine-tuned model for image generation. This demonstration facilitates practical comparisons in the interior space visualization process through the generation of design alternatives, as proposed in this study. Additionally, it allows for the observation of the potential for learning various customized design styles.

Table 7: Image Generation using Default and Fine-Tuned Models

Design Style		AI-based generated Images	
Modern			
Contemporary			

Industrial	
Scandinavian	
Bohemian	
Rustic	
Hygge	
Maximalist	
Shabby	
Provence	
Art Nouveau	
Oriental	
Colonial	
Zen	

Junk			

5. CONCLUSION

In this study, we investigated a method for fine-tuning design styles using Image-Gen AI and automatically generating spatial images reflecting interior design representation. During the research process, we conducted an evaluation of the recognition level for interior design styles by the base model. We implemented a design style visualization model based on detailed keywords for the Brutalism style, which was chosen as one of the fine-tuning targets. The model effectively learned the characteristics of the style and demonstrated the ability to intricately represent the visual attributes of the style.

Through comparative analysis with the base model, we confirmed the high likelihood of visualizing the features of the style, thus validating the capability to effectively visualize spaces that align with user preferences through additional fine-tuning for interior design styles. Furthermore, this research approach showcases the potential for Image-Gen AI to be utilized in various fields, and we aim to suggest its applicability in future research and application domains.

6. ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MIST) (No. NRF-2021R1A4A1032306)

REFERENCES

Ah-Soon, C., & Tombre, K. (1997). Variations on the analysis of architectural drawings. In Proceedings of the fourth international conference on document analysis and recognition (Vol. 1, pp. 347-351). IEEE.

Brock, A., Donahue, J., & Simonyan, K. (2018). Large scale GAN training for high fidelity natural image synthesis. arXiv preprint arXiv:1809.11096.

Ching, Francis DK. (2011). A visual dictionary of architecture: John Wiley & Sons.

Chiu, M. L. (1995). Collaborative design in CAAD studios: shared ideas, resources, and representations. In Proceedings of International Conference on CAAD Future Vol. 95, (pp. 749-759).

Eckert, C.; Stacey, M. Sources of inspiration: A language of design. Des. Stud. 2000, 21, 523–538.

Goldschmidt, G. Creative architectural design: Reference versus precedence. J. Archit. Plan. Res. 1998, 258–270.

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. Advances in neural information processing systems, 33, 6840-6851.

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1125-1134).

Jeong, H., & Lee, J. K. (2023). Study on the Applicability of Image Generation AI in Interior Architecture: Generating Images Based on Interior Design Styles. Proceeding of Spring Annual Conference of KHA, Vol.35, No.1 (pp. 247-250), Jeju, Korea.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of stylegan. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 8110-8119).

Kim, J., & Lee, J.K. (2020) Stochastic Detection of Interior Design Styles Using a Deep-Learning Model for Reference Images. Appl. Sci. 10, 7299. https://doi.org/10.3390/app10207299.

Kim, J., Song, J., & Lee, J. K. (2019, January). Approach to auto-recognition of design elements for the intelligent management of interior pictures. In Proceedings of the 24th International Conference on Computer-Aided Architectural Design Research in Asia: Intelligent and Informed, CAADRIA (pp. 785-794).

Lee, S.H, Kim, J.S, Song, J.Y & Lee, J.K. (2020). Augmented Reality-Based Approach for Design On-Site Visualization - Focusing on the Example of Space Improvement Design at a University Academic Information Center. Journal of the Korean Society of Interior Design, 29(1), 97-104.

Lee, J.K., Lee, S., Kim, Y., & Kim, S. (2023). Augmented virtual reality and 360 spatial visualization for supporting user-engaged design, Journal of Computational Design and Engineering, Volume 10(3), Pages 1047–1059, https://doi.org/10.1093/jcde/qwad035.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).

Nichol, A., Dhariwal, P., Ramesh, A., Shyam, P., Mishkin, P., McGrew, B., & Chen, M. (2021). Glide: Towards photorealistic image generation and editing with text-guided diffusion models.

Oppenlaender, J. 2022. The creativity of text-to-image generation. In Proceedings of the 25th International Academic Mindtrek Conference. pp. 192-202.

Oxman, R. (2006). Theory and design in the first digital age. Design studies, 27(3), 229-265.

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. 2022. Hierarchical text-conditional image generation with clip latents. arXiv preprint arXiv:2204.06125.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. 2022. High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684-10695.

Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E. L., & Norouzi, M. 2022. Photorealistic text-to-image diffusion models with deep language understanding. Advances in Neural Information Processing Systems, 35, 36479-36494.

Song, J., Meng, C., & Ermon, S. (2020). Denoising diffusion implicit models. arXiv preprint arXiv:2010.02502.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & Polosukhin, I. 2017. Attention is all you need. Advances in neural information processing systems, 30.