Advanced Engineering Informatics 62 (2024) 102686

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Creating spatial visualizations using fine-tuned interior design style models informed by user preferences

Jin-Kook Lee [a], Hyun Jeong [a], Youngchae Kim [a], Seung Hyun Cha [b,*]

[a] Department of Interior Architecture and Built Environment, Yonsei University, Seoul 03722, Republic of Korea
[b] Graduate School of Culture Technology, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea

ARTICLE INFO

Keywords:
Architectural Design
Interior Design
Spatial Visualization
Image-Generation ai
Model Fine-tuning

ABSTRACT

This study examines the automated creation of spatial visualizations for interior design, emphasizing user preferences over precision. Recognizing design as a reflection of personal identity, we utilize domain-specific, image-fine-tuned AI models to capture the qualitative aspects of various design styles. In interior architecture, design styles are often categorized by shared visual features—like material use, color combinations, and furniture arrangement—based on tacit consensus rather than explicit data. These features significantly impact both the aesthetic and functional aspects of spaces, influenced by historical, cultural, and personal factors. We advanced the field with a text-to-image model that translates descriptive text into visual representations. An extensive evaluation of the default model was conducted, generating over 15,000 images across 25 design styles, which informed the subsequent integration of detailed design knowledge into the model's training. The refinement process included data preparation, textual alignment with image content, and hyperparameter optimization to develop fine-tuned models. Implemented across multiple scenarios, this approach proved successful in combining the nuanced models with the default, creating images that align with user-defined styles. This methodology serves as a tool for generating spatial visualizations that align with user requirements, providing a range of styles that cater to diverse preferences. It highlights the potential of AI in enhancing design visualization and the shift towards personalized, user-centric design solutions.

1. Introduction

Spatial visualization is vital for interpreting and visually expressing the arrangement, colors, materials, lighting, and additional elements within an interior space. Traditionally, this visualization process relied on hand-drawn sketches and two dimensional (2D) renderings, thus limiting the capacity to convey the full depth and detail of interior designs. The advent of sophisticated technology and computer-aided design tools indicated a significant evolution, enabling designers to capture spaces with remarkable precision and detailing every aspect from textures and materials to lighting effects. Recent advancements in high-performance rendering coupled with the rise of image-generation artificial intelligence (image-gen AI) technologies based on large language models (LLMs) have further revolutionized this field. These technologies enable the generation of spatial visualizations that are highly detailed and closely mimic the realism of photographs, thus offering a new paradigm regarding the conceptualization and visualization of interior spaces [50,57,59]. Our study presents a method that harnesses the power of image-gen AI to generate spatial visualizations tailored to individual preferences, informed by a diverse array of interior design styles.

The design style's influence on the aesthetic and functional attributes of space is undeniable. It is informed by a complex interplay of cultural, regional, historical, and personal influences, affecting everything from spatial layout to the selection of materials, color schemes, and the arrangement of furniture and décor [35,66,11,12,52]. These styles reflect distinct design philosophies and aesthetic values and are deeply intertwined with personal identity and lifestyle choices, emphasizing the importance of personalization in spatial visualization [21,16]. However, individual interpretations of design styles can vary widely, leading to the need for a versatile approach that can accommodate a broad spectrum of preferences [5].

This study introduces a methodology for fine-tuning image-gen AI models with keywords that encapsulate specific design styles, enabling

* Corresponding author at: Graduate School of Culture Technology, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea.
E-mail address: shcha@kaist.ac.kr (S.H. Cha).

https://doi.org/10.1016/j.aei.2024.102686
Received 31 October 2023; Received in revised form 24 February 2024; Accepted 1 July 2024
Available online 4 July 2024
1474-0346/

the generation of spatial visualizations that reflect diverse personal tastes and preferences as shown in Fig. 1. By leveraging this approach, we expand the diversity of design styles accessible for spatial visualization, thereby enhancing interior design's personalization. Moreover, this study highlights the potential of image-gen AI in transforming design visualization practices, emphasizing the shift toward more user-centric design and customization in the field of interior architecture.

2. Background

2.1. Spatial design visualization Automation

The evolution of technology is changing how spaces are conceived and realized in the field of interior architecture. One of these changes, spatial visualization, an integral component of interior architectural design, has undergone significant development over the years. Conventionally, designers relied on manual methods, including hand-drawn sketches and 2D drawings, to communicate their ideas for interior spaces [53],[13]. However, with the introduction of advanced technology and computer-aided tools, achieving increasingly realistic and detailed spatial visualization has become possible [33].

The integration of advanced software, computer-aided design, and three-dimensional (3D) modeling tools has empowered designers to create immersive and realistic visualizations of interior spaces [69,62]. This evolution has enabled designers to accurately represent the physical layout and intricate details such as lighting, materials, textures, and even the interplay of shadows. Thus, stakeholders, including clients and project collaborators, can review the proposed design process and experience a lifelike representation of the design before the construction [4,9]; La, 2014; [30,51,65,71]. Spatial visualization has become a fundamental aspect of the interior architecture design process.

Automation techniques fueled by AI and machine learning have further improved the capabilities of spatial visualization [15]. These technologies can analyze design elements, user preferences, and historical data to generate multiple design options, thereby improving the decision-making process [53,72]. Moreover, the evolution of spatial visualization has revolutionized the design process and elevated the user experience. Thus, clients and end-users can now engage with interactive virtual tours, enabling them to navigate through spaces, evaluate design choices, and offer feedback [58]. This level of engagement fosters a deeper understanding of the design intent and facilitates improved communication among stakeholders [14].

With technological advancement, we can anticipate the emergence of even more sophisticated tools, including real-time rendering, augmented reality (AR), and virtual reality (VR) experiences [76,49,61,54]. These innovations will further bridge the gap between the virtual and physical realms, offering designers and clients a seamless and immersive design exploration process [1,39]. From manual sketches to immersive virtual experiences, this technology has redefined how interior spaces are visualized, communicated, and ultimately realized. Thus, with the advancement of automated technologies powered by AI and machine learning, spatial visualization will remain integral in creating harmonious, functional, and visually captivating indoor environments [44]. Accordingly, in this study, we use image-gen AI to construct an AI model that generates interior space images reflecting various interior styles and offer practical examples of its application.

2.2. Machine learning for spatial Visualization: Applying interior design styles

Machine learning is a transformative technology that has made significant strides across various domains, revolutionizing how we approach complex tasks. Image training is a prominent application of machine learning, wherein algorithms are designed to recognize and interpret visual data with remarkable accuracy [43].

In image training, the use of neural networks has proven a game-changer. Inspired by the structure of the human brain, neural networks are composed of interconnected layers that process and extract features from data [40,28]. Convolutional neural networks (CNNs), a subset of neural networks, excel in image-related tasks because of their ability to automatically learn hierarchical features from images [70,22,43]. Furthermore, the image training process involves feeding large datasets into these neural networks, allowing them to learn patterns, shapes, and relationships in the data. Thus, through iterative training and optimization, the networks gradually refine their parameters to improve accuracy and predictive power [31,20].

Generative adversarial networks (GANs) have also emerged as a pivotal model in image training. GANs comprise two networks, namely, a generator and a discriminator, which engage in a competitive process [77,55,73,26,42]. The generator produces images, and the discriminator evaluates their authenticity. Through this adversarial training, GANs can generate highly realistic images that are often indistinguishable from actual photographs [41,3]. Image training models have been applied in the architectural and interior design fields. Moreover, efforts have been made to employ these techniques to identify the information represented in interior design styles and make it comprehensible to computers [25,38,8]. Kim et al. [2019] and Kim and Lee [36] proposed a method that uses reference images and a deep learning model to probabilistically detect and incorporate interior design style information. Furthermore, Liu et al. [45] used deep learning to explore trends in interior design in different regions. Furthermore, Kim and Lee [37] performed the automatic recognition of room usage within indoor photographs of South Korean apartments, which was integrated into an intelligent management system for interior reference images. Through

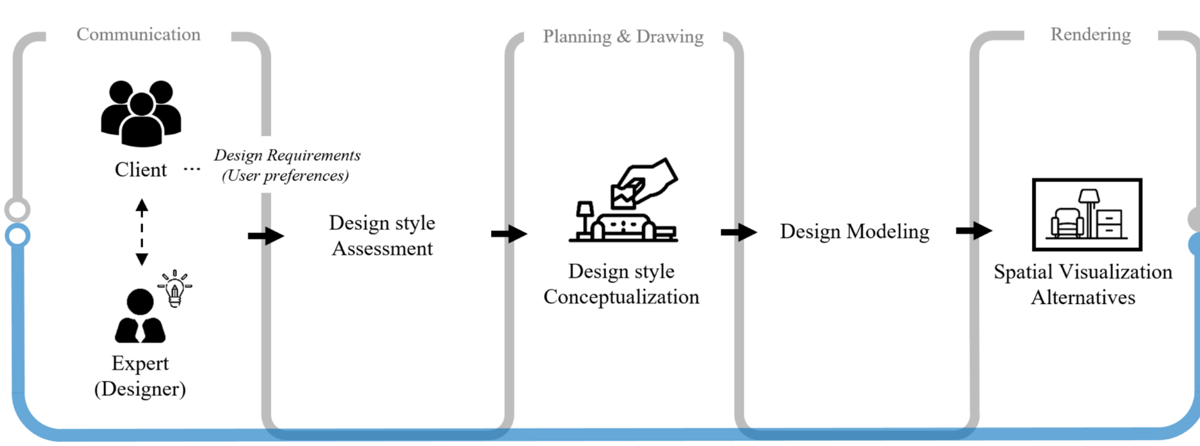


Fig. 1. Overview of the proposed approach.

these investigations, the application of machine learning models for the enhancement of spatial design visualization has been examined consistently, with ongoing endeavors aimed at improving quality outcomes as shown in Fig. 2.

In summary, various approaches have been proposed to automatically recognize interior spaces and components using machine learning. Although data-driven design recognition models have been presented, limited discussion exists regarding the practical implementation of spatial design visualization. However, with the advancement of image generation techniques, image-trained models are becoming increasingly sophisticated. For example, diffusion-based methods allow models trained on one dataset to be fine-tuned for other relevant datasets, thereby reducing the need for extensive training on diverse design recognition [47,29]. Therefore, this study proposes a machine learning-based approach to spatial visualization that can efficiently generate interior design styles.

2.3. Potentials for automating spatial visualization using image-gen AI

Considerable studies explore the field of deep learning-powered image synthesis [34,60,39]. However, the full potential for effectively visualizing spatial designs remains to be thoroughly explored. This study acknowledges the potential impact of image-gen AI in interior architecture and proposes a methodology aimed at achieving effective spatial visualization via the integration of design styles. The methodology harnesses the capabilities of advanced machine learning techniques, leveraging the progress made in large language models (LLMs) and image generation. Consequently, the once-distant prospect of generating spatial visualization images based on textual descriptions is well within reach currently.

Text-to-image (txt2img) generation produces highly realistic visual representations [18,63,17]. Hence, it emerges as a versatile tool with the potential to create diverse spatial visualization content. As AI technology continues to evolve, the translation of textual concepts into tangible visual outputs can play a pivotal role within interior architecture [64]. This evolution extends beyond conventional approaches, unlocking new avenues for innovative and imaginative possibilities in image-gen AI. Furthermore, it is poised to become an indispensable tool that bridges the gap between text-based ideas and realistic visual representations in interior architecture. This integration aligns with the broader trend of AI enhancing creative processes across various industries [46].

3. Intensive test on image-gen AI for design styles

3.1. Image-gen AI models

In the evolving landscape of AI-based generative models, txt2img generation, in particular, has progressed significantly [10]; Oppenlaender, 2023. Entities such as models "D," "M" [48], and "SD" are noteworthy. These models are distinguished in the field of image-gen AI, each marked by unique technological attributes, strengths, and limitations [Oppenlaender, 2023. This study conducts a comprehensive examination and comparative analysis of the txt2img generation capabilities inherent in these models. Moreover, it deliberately excludes considerations related to the user interface and usability. This study explores the technical intricacies underlying the txt2img generation functionalities offered by these models.

Model "D' is renowned for its ability to comprehend and interpret textual prompts to create detailed and intricate images [75,56]. However, its complexity and resource-intensive nature may yield longer processing times. Furthermore, owing to its reliance on textual inputs, generating images without textual cues could be challenging. Model "M' is another notable image-gen AI model that specializes in producing images inspired by a specific visual style or artistic concepts [50]. It emphasizes artistic interpretation and style transfer, enabling users to generate images closely aligning with the desired aesthetics. However, its focus on artistic expression might yield less accurate representations of real-world objects or scenes. In addition, Model "SD' is an image-gen AI model that prioritizes stability and accuracy in image generation [Rombach, 2022. Using small large language models (sLLMs), it generates images with well-defined attributes and features. This approach ensures that the generated images precisely match the intended design specifications. The model's strength is its ability to generate realistic and detailed images, making it suitable for scenarios where precision is paramount [67,6].

Despite its strengths, it also has specific limitations. Users should consider their projects and the desired outcomes when selecting an image-gen AI model. Thus, the study suggests that model SD is particularly well-suited for fine-tuning models focused on interior design styles due to its emphasis on stability, accuracy, and generation of images with well-defined attributes. Precision and faithful representation of design elements are crucial in interior design. Model SD's precision, attention to detail, realistic visuals, control over attributes, design consistency, and suitability for iterative design processes make it a compelling choice for fine-tuning models in this domain. Furthermore, its capabilities align perfectly with the need for precision and visual

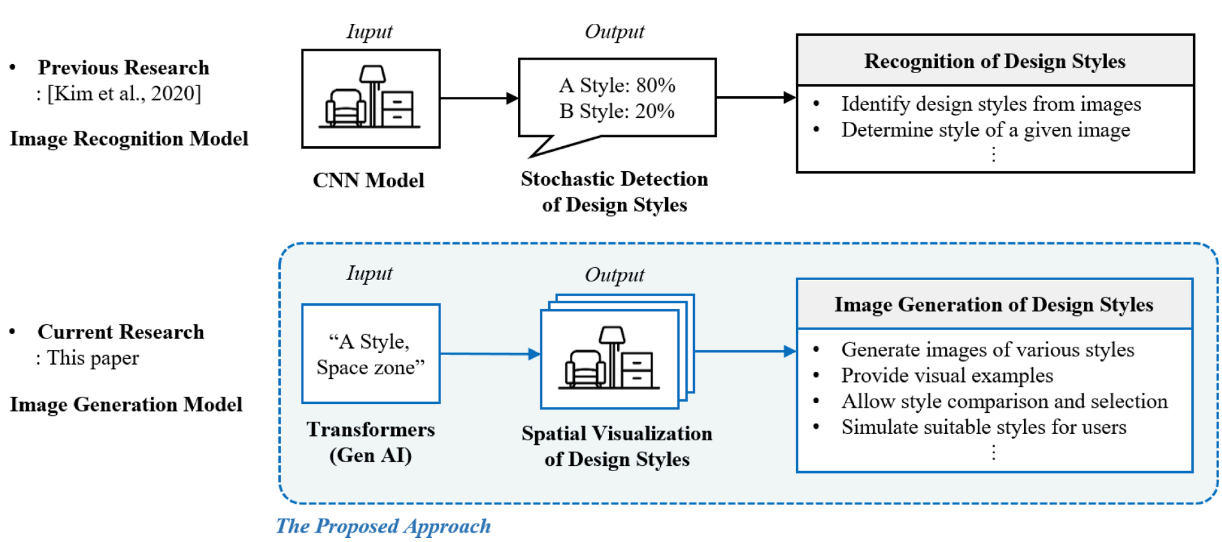


Fig. 2. Comparison between previous CNN-based image recognition model for interior design styles [Kim et al., 2020] and our proposed generation approach.

Table 1
Standardized template of prompts for design styles.

Prompt	Positive	Negative
SDP	Target design style (e.g., Modern), spatial zoning (e.g., a living room), floor-to-ceiling windows	Excessive clutter and chaos, disorganized layout, dull and uninspiring colors, overwhelming patterns and textures, inadequate lighting and ambiance
VQP	Professional photograph, photorealistic rendering, enhanced detail, V-ray rendering, full HD, masterpiece, highly detailed, high quality, 8 K, full shot, high depth of field (f/22, 35 mm), high-key lighting, realistic shadows, and attention to detail	Bad proportion, low quality, awkward shadows, pixelated textures, worst, noisy, unrealistic reflections, normal quality, watermark, bad perspective, confusing details, blurry textured, faint, text, tacky, crowded, and signature

fidelity necessary to effectively visualize and communicate diverse interior design aesthetics. Moreover, as an open-source model, SD offers an adaptable usage for researchers. The transparency of the source code enables researchers to understand and modify the underlying algorithms, enabling tailored applications and contributing to advancements in the field of image generation.

3.2. Image generation test for design styles

In this section, we conducted an intensive test to evaluate the recognition level of 25 distinct design styles using the image-gen AI Model “SD’. The selection of design styles prioritized diversity to align with the objectives of the comprehension evaluation test. The criteria for selection were as follows: 1) diversity (cultural, regional, and temporal diversity, 2) comparability (the presence of visually distinctive characteristics allowing for comparative analysis between styles), and 3) popularity and recognition. Based on these criteria and considering their distinctive features and similarities, we selected the following 25 design styles [52,68,19,7]: Modern, Contemporary, Rustic, Scandinavian, Industrial, Hygge, Sustainable, Retro, Zen, Brutalism, Mediterranean, Oriental, Shabby, Provence, Cottage core, Victorian, Tudor, French, Neo-classic, Bohemian, Art Nouveau, Maximalist, Kinfolk, Eclectic, and Junk. The image generation test, based on the 25 selected design styles, was conducted using the txt2img approach. The operations involved in generating images using image-gen AI are as follows:

$$genTarget\,(\text{"target"},\ AIM,\ Param) \rightarrow GI \tag{1}$$

The *genTarget()* fun‘ction is pivotal in the image generation process, considering “target” (the designated design style), *AIM* (the default image-gen AI model), and *Param* (parameters including scene description prompt (*SDP*) and visual quality prompt (*VQP*)) as inputs to produce generated images (*GI*). In addition, *SDP* is crucial for evaluating how well the generated images understand the design style. To this end, design style-related prompts were input concisely to provide minimal information about the design style and spatial zoning (e.g., a living room), preventing the default model from making inferences based on textual descriptions. Meanwhile, *VQP* guides the model to meet high standards of visual quality during image generation, and it provides criteria for generating visually appealing, realistic, and effectively composed images. Accordingly, the text prompts were categorized into both reflective and nonreflective aspects, classifying them into positive and negative aspects, as structured in Table 1.

$$if\ (Most\ GI_i \notin TI)\ then \tag{2}$$

$$\text{RecognitionLevel}\,(GI,\ Keywords)\ =\ \frac{1}{\text{N}}\sum\nolimits_{i=1}^{N} 1\,(GI_i\ reflects\ Keywords) \tag{2.1}$$

$$if\ \text{RecognitionLevel}\,(GI,\ Keywords)\ <\ \text{Threshold} \Rightarrow \text{Proceed to Fine}-\text{Tuning} \tag{2.2}$$

Operation (2) is instrumental in recognizing when most generated images (GI_i) do not align with the target image (*TI*), serving as a mechanism to identify mismatches between *GI* s and *TI*. Moreover, Operation (2.1) quantifies the extent to which generated images incorporate key elements of the target design style, with 1 being an indicator function that assesses the presence of these elements within each GI_i. A low recognition level, as determined by Operation (2.2), triggers the fine-tuning process, suggesting that the AI’s default output does not sufficiently capture the target design’s characteristics.

$$getFTM(target \vee TI) \rightarrow FTM \tag{3}$$

Following a low recognition level, Operation (3) outlines the transition from the *AIM* to a FTM, incorporating both the target design style and the *TI* to enhance model accuracy.

$$genTarget(target, FTM, Param) \rightarrow TI \tag{4}$$

Operation (4) depicts image generation using the *FTM*, aiming to produce images that more accurately reflect the *TI*, thereby improving the model’s fidelity to the designated design style. The recognition mechanism in the intensive test, primarily through Operation (2) and its derivatives (2.1 and 2.2), plays a vital role in assessing the AI model’s performance in generating design-congruent images. By evaluating the incorporation of key design elements into *GIs*, this mechanism facilitates the identification and rectification of instances where the model’s output falls short of expectations. Therefore, these operations and functions are integral to our methodology, ensuring the generation of images that adhere closely to the desired design style, with the flexibility to adjust parameters based on specific design styles and study objectives.

3.3. Model fine-tuning for expanded use scenarios

According to the previous discussion regarding the image generation process, we evaluate the recognition level associated with 25 design styles using the generated images. Precisely defining such design styles are challenging because of their qualitative nature, and interpretations may vary depending on individual perspectives. However, they are primarily derived from common visual characteristics that frequently appear in various design cases. These visual characteristics can be expressed using style keywords that encompass design elements, finishes, colors, and more. Therefore, we evaluated the default model’s understanding of 25 interior design styles based on four criteria: 1) understanding of style-related keywords (i.e., color and ambiance), 2) expression of style-related material finishes, 3) composition and furniture arrangement within the target space (living room), and 4) image quality level (i.e., resolution, aspect ratio, and form). Furthermore, the default model consistently demonstrated stable generation capabilities and produced high-quality images for most styles. However, it exhibited limited comprehension for styles with lower recognition or those reflecting current trends, such as “Brutalism, Zen” or “Retro, Sustainable,” resulting in restricted expressive abilities for these specific styles.

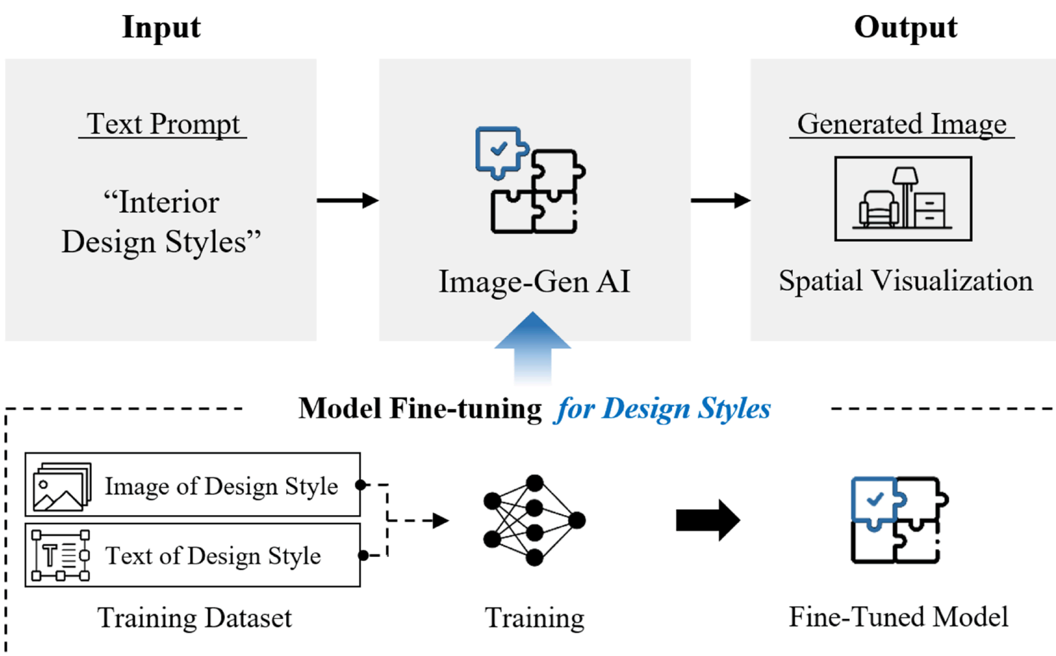


Fig. 3. Proposed approach to model fine-tuning for image-gen AI-assisted spatial visualization.

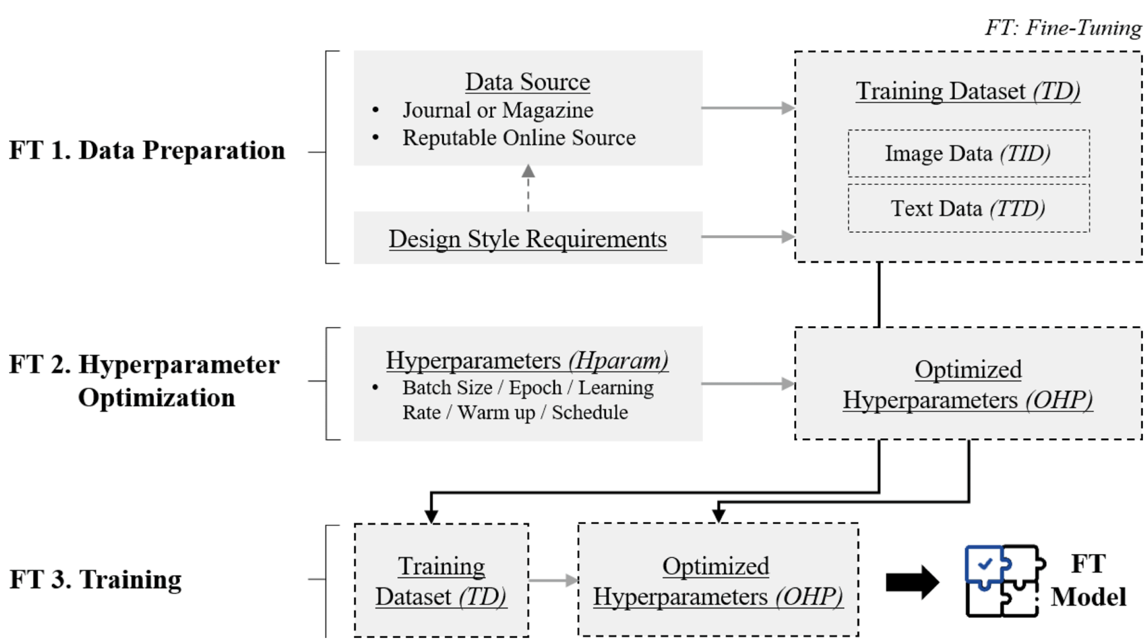


Fig. 4. Overall process of model fine-tuning.

This limitation is attributed to the nature of AI models that generate results based on the provided training data. Model fine-tuning was conducted to address these constraints and enhance the accuracy of image generation for specific design styles as shown in Fig. 3.

4. Model fine-tuning of design styles

4.1. Overall process

In this section, the model fine-tuning process (FT) involves the following three steps as shown in Fig. 4. First, FT1) data preparation (including preprocessing) entails preparing and preprocessing the data specific to the target. This includes tasks such as collecting reference images, cleaning and organizing the dataset, and ensuring data format consistency. Second, FT2) hyperparameter optimization centers around optimizing the model's hyperparameters to enhance its performance in generating target images. This optimization involves fine-tuning parameters such as learning rate, batch size, and network architecture to achieve improved results. Finally, FT3) training involves training the model using the prepared dataset and optimized hyperparameters. The model grasps the specific characteristics and nuances of the target design style, enabling it to generate images that closely align with the desired style. Following these three procedures, model fine-tuning was conducted on the default model to enhance its proficiency in generating images of the target design style.

4.2. Data preparation

To train Design Style *DS*, we require training data $TD = \{TID, TTD\}$, which consists of pairs of images and corresponding text, serving as design style references. Training Data(*TD*) for fine-tuning consists of two components: training image data (*TID*) and training text data (*TTD*) associated with design styles.

To obtain superior quality and consistent representations in the *TID*, we carefully extracted image data from specialized interior design magazines and reputable online sources, which is one of the methods for preparing such data. The criteria for collecting image data include the following aspects: 1) inclusion of images that capture indoor spaces from floor to ceiling; 2) categorization of images into well-defined stylistic categories; 3) presence of keyword elements within images that represent specific design styles. Subsequently, preprocessing tasks such as image resizing were performed. The preprocessed design style images serve as training image data. We conducted a comprehensive captioning effort by adding textual annotations to the gathered images to enhance the precision of each distinctive style. Each training image data is mapped one-to-one with the training text data.

4.3. Hyperparameter optimization

This study applied the low-rank adaptation (LoRA) methodology for fine-tuning models It is particularly useful in natural language processing and computer vision and is an efficient method for fine-tuning large pretrained models [24]. This is achieved by introducing small, trainable matrices that interact with the original model's fixed weights [32]. A major advantage of LoRA is its ability to fine-tune large models more efficiently for specific tasks or datasets. Moreover, the original weights of the model remain unchanged, preserving the model's general knowledge and capabilities. The only components trained are the added low-rank matrices, which require significantly less computational power and resources than training the entire model [74]. Using LoRA in this study, we could tailor the model for specific purposes, avoiding the intensive resource demands typically associated with training large-scale AI models.

The efficiency of LoRA can be enhanced via hyperparameter optimization. Although hyperparameter optimization encompasses various settings, this study primarily focused on optimizing parameters such as the train batch size, epochs, learning rate, learning rate scheduler, and learning rate warmup. The goal of this optimization is to strike a balance where the low-rank matrices are adequate for the model to adapt to new tasks or datasets without imposing an excessive burden on computational resources. Thus, employing this combined approach can enhance the performance of fine-tuned model while leveraging the advantages of using pretrained models.

In this study, we adjusted the hyperparameters and fine-tuned each combination. Subsequently, we extracted the resulting loss values to assess the image quality. Our analysis focused on the effects of various hyperparameter configurations to primarily identify those that yield the best performance in the effective generation of interior design style images. The *HParam* and operator *optimizeHyperparameter*() can be outlined as follows:

$$HParam = \{hparam_j \middle| 1 < j \leq n; n = number\ of\ hyperparameters\ for\ training\ AIM\}$$

$$optimizeHyperparameter(HParam) \rightarrow OHP$$

4.4. Training

During the training phase, using the training dataset established in Section 4.2, fine-tuning is performed. The fine-tuned LoRA model is employed alongside the default model, and the application rate is denoted by weight *W*. To ensure the training process's versatility, this study introduces the *FineTune()* operator. Thus, the fine-tuned model can be used to generate *TIs* in the following operations:

$$FineTune(AIM, TID, TTD, OHP) \rightarrow FTM$$

$$genTarget(FTM \vee (FTM, W), SDP, VQP) \rightarrow TI$$

4.5. Test and evaluation of fine-tuned models and their application

This integration enabled us to train the model using the gathered dataset, including both reference images and textual annotations. Through the training process, the model learned and improved its image generation capabilities, resulting in high-quality images that effectively capture the essence of the designated design styles. Moreover, by using the refined dataset and optimized hyperparameters, we effectively facilitated the generation of design visualization images that are more precise and visually captivating. In this section, we proceed with the testing, evaluation, and application of the fine-tuning models.

The LoRA-based fine-tuning used in this study employed the mean squared error (MSE) loss function, which was selected to minimize the difference between the generated and training dataset (original) images, considering minimizing pixel values. Hyperparameters encompass various settings; however, this study focused on optimizing parameters such as the train batch size, epochs, learning rate, learning rate scheduler, and learning rate warmup. These parameters were optimized to enhance the fine-tuning model performance. Thus, the final quality of the fine-tuned model was evaluated based on various criteria, one of which was the loss value that represents the difference between the model's predictions and the actual results. The final loss values for the superior and inferior models were 0.02 and 0.03, respectively, indicating a 33.3 % difference. This difference in convergence levels between the two cases suggests that the superior model had a lower loss value than the inferior model, implying that the superior model exhibited better performance. Furthermore, these results demonstrate that the lower loss value of the superior model indicates superior generalization ability for the dataset.

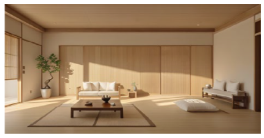
Accordingly, the weight parameters of the selected fine-tuned model determine the extent to which the model contributes to the final output. For example, a weight of 0.1 signifies that the fine-tuned model contributes 10 % to the generated outcome. To assess this approach's effectiveness, we conducted tests using the same prompt while varying the LoRA weight. Table 2 presents the outcomes of the tests, displaying the generated images under different LoRA weight configurations. The results reveal that the fine-tuned model noticeably influences the generation of design style images that were previously unattainable. Higher LoRA weights correspond to a more significant effect of the fine-tuned model on the image generation process.

5. Demonstration

5.1. Overview

Interior design encompasses various styles and trends based on the combination of design elements, color choices, and materials used. Furthermore, each space reflects unique preferences and styles influenced by culture, region, and individual tastes [66]. This diversity facilitates the emergence of new design styles and various adaptations. However, the default model of image-gen AI may not always accurately capture the diversity of design styles and stay up-to-date with the latest trends. Hence, the fine-tuning approach proposed in this study enables the effective generation of customized designs tailored to user preferences. In addition, style keywords play a crucial role in categorizing design styles, facilitating the visual representation of various styles through keyword combinations. In this section, we have constructed design style models that encompass unrecognized design styles and the

Table 2
Test of fine-tuned models according to weight *W*.

W	**0.3**	**0.6**	**0.9**
GI			

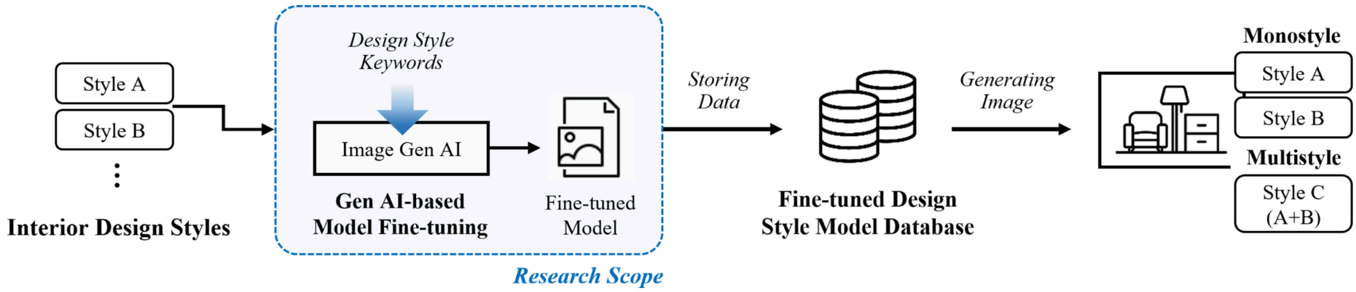


Fig. 5. Demonstration overview: 25 interior design style generation using base model and fine-tuned models.

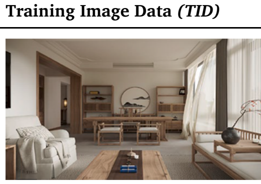
latest trend styles as shown in Fig. 5. We employed these models to generate images showcasing 25 diverse design styles, including monostyles and multistyles (mix and match). Thus, we demonstrate the ability to generate spatial visualization images for specific design styles that were challenging for the default model. Consequently, we effectively visualize several design styles based on trends and user preferences, enabling the creation of spaces that align with user objectives.

5.2. Model fine-tuning of design styles

In this section, we describe the procedures performed to fine-tune the default model for a specific design style. These procedures include FT1) data preparation, FT2) hyperparameter optimization, and FT3) training. During the data preparation stage, as detailed in Section 4.2, the training data for fine-tuning are divided into two categories, namely, reference images and textual annotations offering descriptive information about each image. Our study facilitated the comparison of visualizations of various design styles by standardizing the target space. Consequently, we focused on the living room, a central area in residential spaces, as our designated target space.

For a high-quality dataset, we meticulously collected image data from specialized interior design magazines, such as Architectural Digest and Interior Design, and from reputable online sources such as Houzz [Architectural [2]; Interior [27,23]. The criteria for image data collection are as follows: 1) inclusion of images that comprehensively capture indoor spaces from floor to ceiling, 2) organization of images into clearly defined stylistic categories, and 3) presence of keyword elements within images that signify specific design styles. Subsequently, all the collected images underwent preprocessing to prepare the dataset. This preprocessing included resizing the images to the most effective size for LoRA training, which is 512 x 512 pixels. Subsequently, we performed a labeling task by manually adding text annotations to the preprocessed images to represent the design style information for each image. To enhance the accuracy, prompt engineering was conducted for each design style. It involves designing and fine-tuning the text prompts that are input into the model when using image-gen AI. This process guides the model to perform specific tasks and generate outputs in the desired direction. Moreover, the design styles under investigation in this study are subjective in nature, and their interpretation can vary based on individual perspectives, making it challenging to provide a precise definition. Consequently, for prompt engineering in this study, we extracted detailed keywords related to design styles from the collected images, focusing on design elements, finishing materials, colors, and mood. Examples of these keywords utilized in actual fine-tuning are provided in Table 3. Subsequently, we constructed approximately 400 training datasets by one-to-one mapping of the processed images with textual data. Fig. 6 presents an overview of the FT1 step.

Table 3
An example training dataset. We constructed a text prompt using detailed descriptions of *content* (e.g., a Space Zoning), *style* (e.g., a design style), and *scene*.

	Training Image Data *(TID)*	**Training Text Data *(TTD)***
TD		A *Zen-style* interior in *a living room* characterized by minimalism, serenity, balance, and Japanese aesthetics. Utilize finishing materials such as natural wood, soft textiles, and subdued hues to create a serene ambiance, harmonizing with the neutral tones and color scheme to establish a truly calming atmosphere.

The fine-tuned model's quality is significantly influenced by the optimization of hyperparameters such as epochs, learning rate, learning rate scheduler, and learning rate warmup. These parameters directly impact how well the model can adapt and learn when incorporating new data. The epoch parameter determines the number of training sessions, while the learning rate parameter controls the size of weight adjustments, striking a balance between how rapidly the model converges and its stability during training. Moreover, the learning rate scheduler adjusts the learning rate during the training process, facilitating efficient training. In addition, the learning rate warmup gradually increases the learning rate at the beginning of training to ensure a stable process. Accurately configuring these hyperparameters is critical for achieving a high-quality model extension within the LoRA framework. By experimenting with different types and configurations of hyperparameters, we conducted various training experiments to achieve superior quality in the target design styles. As per the test results, the hyperparameters for further training were set with a batch size of 2, 150 epochs, and a learning rate of 0.0001, while other parameters were configured as detailed in Table 4. In this demonstration, we set the optimization criterion for hyperparameters to achieve loss values below 0.02 (refer to Section 4.3). The final values for these hyperparameters were determined through trial-and-error optimization and are detailed in Table 4.

During the FT3) training step, the training procedure used training

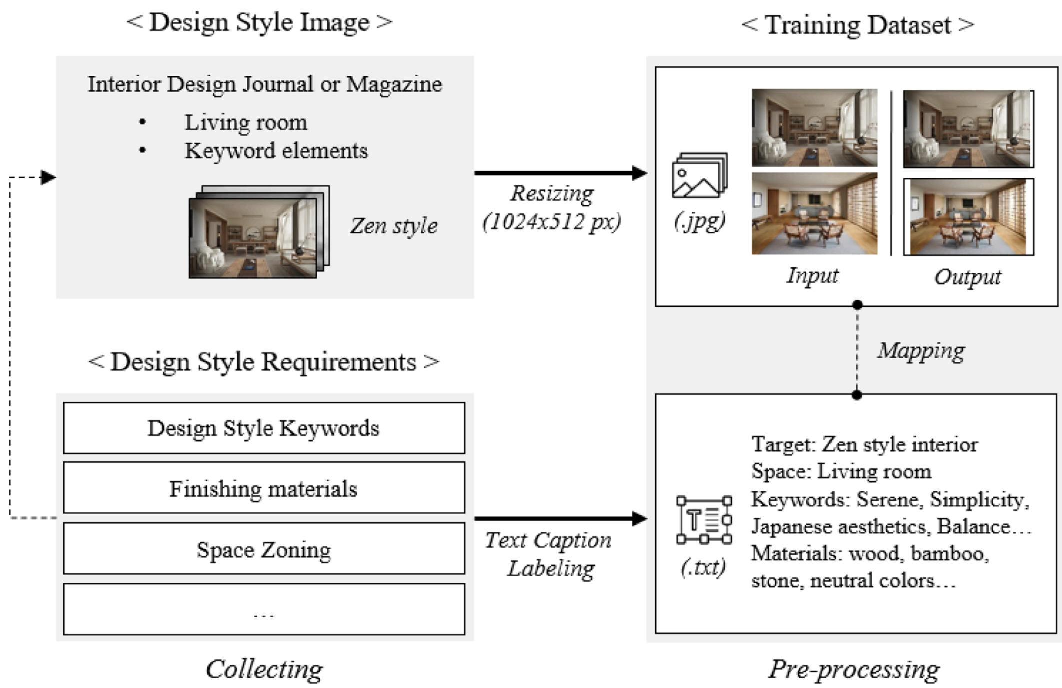


Fig. 6. Overview of FT1: data preparation.

Table 4
FT2: Hyperparameter optimization for model fine-tuning.

Hyperparameters	Purpose	Value
Batch size for training	Optimized to balance efficiency and performance by determining number of training examples processed together in each iteration	2
Epoch (Training steps)	Controls the number of times the model goes through the entire training dataset, striking a balance between performance and computational	150
Learning rate	Determines the step size for parameter updates during training optimized to achieve effective convergence and performance	0.0001
Learning rate scheduler	Specifies how the learning rate changes throughout the training	constant
Learning rate warmup	Gradually increasing the learning rate at the beginning of training to stabilize the optimization process and promote smoother and more stable learning	10

data and specific hyperparameter configurations to create distinct design styles, and it occurred on a local computer equipped with an RTX A6000 GPU with a memory capacity of 47.5 GB. The training procedure was completed within 20 min. Following the fine-tuning of the model, a LoRA model was produced, which had a compact size of 144 MB and was saved in the (.safetensors) format. This refined model, obtained from fine-tuning, is employed alongside the default model for generating images. An overview of the FT3) step is shown in Fig. 7.

The model was trained on the style of *TID* and specific design requirements derived from *TTD*. Furthermore, this dataset was employed to produce interior space images guided by prompts incorporating distinct design styles. Configuration settings applied during the image generation process using Image-Gen AI, including sampling approaches, sampling intervals, dimensions (width and height), and CFG scale, significantly influence the outcomes. Thus, in this section, we have identified optimal configuration settings aligning with the objective of generating interior space images that portray each unique design style. Table 5 presents elaborate information regarding these settings. The two cases presented in Table 6 generate interior space images using the same input. The only difference between Cases A and B is the use of the fine-tuned model developed in this study. The outcomes demonstrate the successful achievement of our objective of generating images that embody design styles previously unrecognized, achieved via the fine-tuning of the image-gen AI model.

5.3. Generating spatial visualization images reflecting various design styles

In this study, the primary objective of fine-tuning the image-gen AI model was to understand diverse trend-sensitive design styles. Interior design encompasses various styles and trends characterized by unique design elements, color palettes, and materials. In addition, each space possesses a distinct style influenced by factors such as region, culture, and user preferences. Nevertheless, the default model may not always accurately capture the latest trends in interior design.

Meanwhile, the approach proposed in this study enables the generation of interior space images reflecting various design styles and facilitates the emergence and evolution of new design styles and variations. Therefore, in this section, we conduct demonstrations following the approach outlined in Section 5.2. Using both the default and fine-tuned models, we generated images representing 25 different types of design styles, and the target space was limited to living rooms. Table 7 displays mono-style case images generated to represent specific design styles. To assess whether these generated results appropriately capture a particular style, we evaluated them based on three criteria, namely, 1) inclusion of design style-related keywords; 2) Spatial arrangement, furniture placement, and the degree of form distortion; 3) Image quality, including resolution and aspect ratio. As this study primarily focuses on design styles, we determined the appropriateness of Criterion 1) by selecting images that met Criteria 2) and 3). Table 7 reveals that the generated images effectively incorporate relevant keywords at a high level for each design style, reaffirming the fine-tuned model's effectiveness in generating design styles while considering user preferences.

By combining various styles and trends and introducing novel design concepts, creativity can be inspired. Our approach involved careful

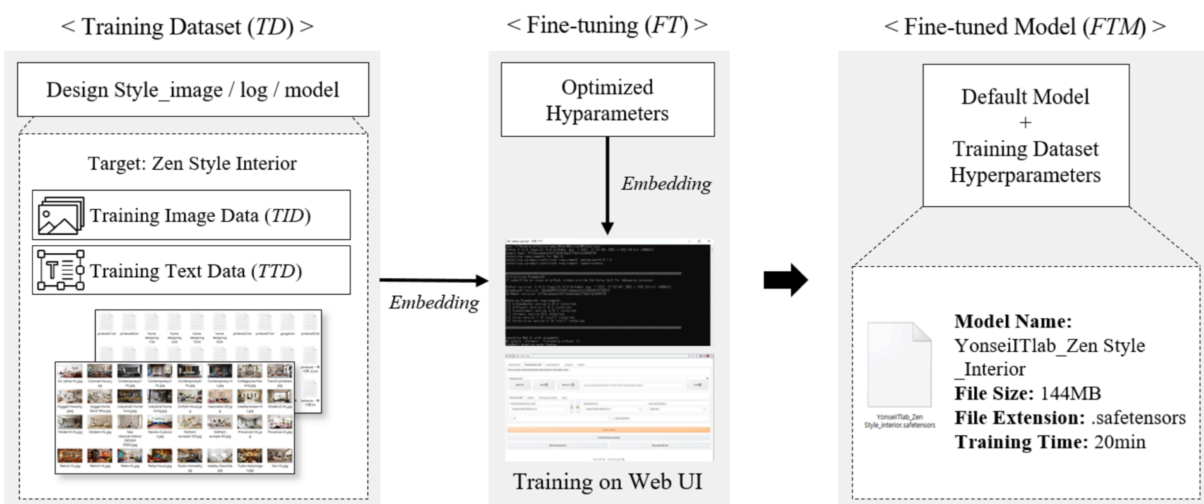


Fig. 7. Overview of FT3: training.

Table 5
Configuration settings for image generation.

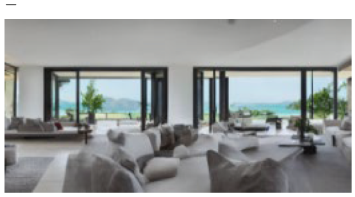
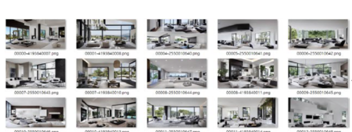
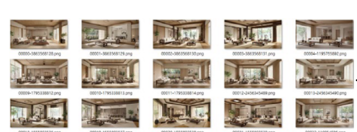
Base model	Sampling method	Sampling steps	CFG scale	Resolution
SD v1.5 ckpt	DPM + 2 M Karras	25	13	1024 × 512 (2:1)

curation and harmonization of elements, including design keywords, material finishes, and color palettes. This deliberate selection aimed to break free from conventional designs and create unique design environments by blending diverse styles. Moreover, by using a fine-tuned model that embraces user preferences and facilitates the mixing and matching of interior design elements, we are empowered to efficiently offer customized design solutions. As previously mentioned, the resulting image dataset serves as a valuable resource in the field of interior design and as a foundation for exploring design styles [38]. Furthermore, the archiving of this dataset serves a dual purpose—as a tool for training other generative AI models and as a source for fostering creative image generation. To visualize the outcomes of our multistyle (mix and match) approach, Table 8 presents the findings of image generation.

5.4. Application of the proposed approach to diverse usage scenarios

In this section, we generate interior space images using the image-to-image generation method, employing a fine-tuned model to reflect interior design styles. This study aimed to enable the visualization of virtual styles solely through images, without additional modeling. This simplifies the intricate process of design visualization, supporting spatial redesign decisions through user experience and comparative analysis. Thus, applying the fine-tuned model to interior space visualization facilitates intuitive and realistic design comparisons. Table 9 presents

Table 6
Qualitative comparison of image generation using fine-tuned models.

			Case A	Case B
Input	Type	Space Zoning	A living room	
		Design Style	Zen Style	
	Prompt	*SDP*	Zen-style interior, a living room. Use clean lines, neutral colors, and natural materials. Include a low-profile sofa, a Japanese-style coffee table, and a small indoor plant to add a touch of nature.	
		VQP	PP_VQP NP_VQP	
	Model	*FTM*	–	YonseiITlab_Zen Style_Interior.safetensors
Output	AI-based generated image (in large)			
	AI-based generated images		 …	 …
	Quick Review		– Specific drawing style: low – Fidelity of domain issue: low – Image quality: low	– Specific design style: high – Fidelity of domain issue: high – Image quality: high

(ex) PP_VQP = "Professional photograph, photorealistic rendering, enhanced detail, v-ray rendering, full HD, masterpiece, highly detailed, high quality, 8 k, full shot, deep depth of field, f/22, 35 mm, high-key lighting, realistic shadows".
(ex) NP_VQP = "Bad proportion, low quality, awkward shadows, pixelated textures, worst, noisy, unrealistic reflections, normal quality, watermark, bad perspective, confusing details, blurry textured, blurry, faint, text, tacky, crowded, signature".

Table 7
Results of mono-style image generation.

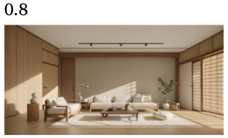
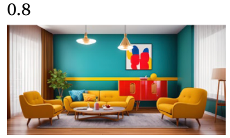
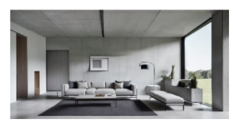
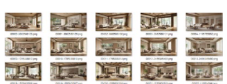
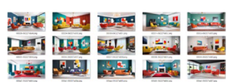
			Case A	**Case B**	**Case C**
Input	Type	Space zoning	A living room		
		Design style	Zen	Retro	Brutalism
	Prompt	*SDP*	Zen-style interior, a living room, Serene, minimalist, neutral, Japanese esthetics, simplicity, natural elements, balance, bamboo, stone	Retro style interior, Living room, vintage, nostalgic, playful patterns, mid-century, wallpapers, vinyl flooring, Formica countertop, Bold colors	Brutalism style interior, living room, unfinished, concrete, minimal, stark, angular, bold, exposed concrete, steel, glass, plywood, metal mesh
		VQP	PP_VQP NP_VQP		
	Model	*FTM*	YonseiITlab_Zen Style_Interior.safetensors	YonseiITlab_Retro Style_Interior.safetensors	YonseiITlab_Brutalism Style_Interior.safetensors
		W	0.8	0.8	0.8
Output	AI-based generated image (in large)				
	AI-based generated images		 …	 …	 …
	Quick review		– Specific design style: high – Fidelity of domain issue: high – Image quality: high		

Table 8
Results of multistyle image generation (mix and match).

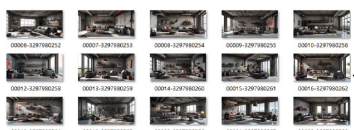
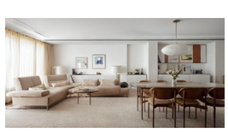
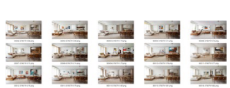
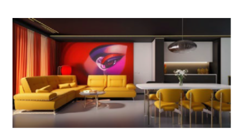
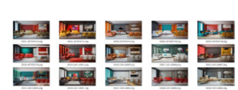
			Case A	Case B
Input	Type	Space Zoning	A living room	
		Design Style	Style A Style B Brutalism Junk	Style A Style B Industrial French
	Prompt	*SDP*	Brutalism style interior, living room, unfinished, concrete, minimalist, stark, angular, bold, exposed concrete, steel, glass, plywood, metal mesh, junk-style interior, living room, eclectics, vintage, repurposed, mismatched, salvaged items, recycled materials, and unique items	Industrial-style interior, a living room, raw, exposed materials and a rugged, urban esthetic brick, concrete, metal, reclaimed wood, French-style interior, a living room, elegant, sophisticated, ornate details, luxurious fabrics, and a sense of refinement
		VQP	PP_VQP NP_VQP	
	Model	*FTM*	YonseiITlab_Brutalism Style_Interior.safetensors YonseiITlab_Junk Style_Interior.safetensors	YonseiITlab_Industrial Style_Interior.safetensors
		W	0.7/0.3	0.5/0.5
Output	AI-based Generated Image (in Large)		Style C (A + B) Brutalism + Junk 	Style C (A + B) Industrial + French 
	AI-based Generated Images		 …	 …
	Quick Review		– Specific design style: high – Fidelity of domain issue: high – Image quality: high	– Specific design style: high – Fidelity of domain issue: high – Image quality: high

Table 9
Results of image-to-image generation.

			Case A	Case B	Case C
Input	Type	Space Zoning	A living room		
		Design Style	Kinfolk	Brutalism	Contemporary + Retro
	Prompt	*SDP*	Kinfolk-style interior, A living room, minimal, organic, natural light, simple lines, serene, light and airy finishes	Brutalism style interior, a living room, unfinished, concrete, minimal, stark, angular, bold, exposed concrete, steel, glass, plywood, metal mesh	Contemporary and retro style interior, a living room, bold and vibrant colors, vintage, nostalgic, playful patterns, mid-century, wallpapers
		VQP	PP_VQP NP_VQP		
	Model	*FTM*	YonseiITlab_Kinfolk Style_Interior.safetensors	YonseiITlab_Brutalism Style_Interior.safetenso	YonseiITlab_Retro Style_Interior.safetensors
	Seed Image				
Output	AI-based Generated Images		  …	  …	  …
	Quick review		– Specific design style: high – Fidelity of domain issue: high – Image quality: high		

examples of images generated by reflecting various styles in the same space.

6. Available resources and limitations of the work

To access a concise preview of the implemented image-gen AI models and downloadable resources, please use the provided webpage hyperlink [Contemporary Design Pictures Archive Web, 2023. The linked page includes shared images depicting 25 distinctive interior design styles, accompanied by design-specific keywords and finishing materials for image generation.This study explored the automatic generation of spatial visualization images representing various design styles using image-gen AI. In this study, we performed intensive testing on the default model and developed fine-tuned models that incorporate diverse design styles. This was performed with consideration for the subjective and varied nature of design styles, aiming to assess the potential of image-gen AI as a tool for spatial visualization that can cater to the preferences of different users.

To this end, we employed prompt engineering to keyword the design elements, finishes, and spatial arrangements associated with each design style, enabling us to build the fine-tuned models. However, this approach was primarily demonstrated in residential living spaces, and its applicability to other space types remains limited. Furthermore, technical challenges remain in accurately representing complex elements such as shadows, reflections, and furniture shapes in the generated images, primarily due to AI models' inherent reliance on training datasets. Consequently, future models can undergo improvement and adjustments to achieve higher levels of accuracy.

In addition, a more generalized approach is required to accommodate the multifaceted nature of design styles, encompassing both emotional and conceptual aspects. Nonetheless, the findings of this study demonstrate that image-gen AI can be a valuable tool in the field of interior architecture, allowing for the intricate implementation of the visual characteristics of design styles, thereby enhances accessibility to visualization without requiring expert assistance. Moreover, while this study primarily focused on the design styles and living rooms, there is a potential for methodological expansion to cater to various requirements based on region, culture, and specific spatial purposes. Further efforts will be required to develop user-friendly applications that consider practical usage into account.

7. Conclusions

This study presents a pioneering exploration of the integration of generative AI into interior architectural design, with a strong emphasis on automating spatial visualization. The combination of AI-generated images and textual descriptions introduces a novel approach for conveying design concepts. Notably, this study emphasizes the revolutionary impact of AI-driven image generation, especially through the txt2img model, acting as a conduit for advanced spatial visualization. This study showcases the efficacy of fine-tuning image-gen AI models using diverse training techniques, enabling the efficient creation of interior architectural visualization images tailored to user preferences. Moreover, the study underscores the potential of AI-powered image generation across various design visualization applications, including the evolution toward user-centric design and personalization trends.

Furthermore, the research explores the role of image-gen AI in enhancing spatial design visualization and bridging the gap between textual descriptions and realistic visual outputs, which aligns with the broader trend of AI transforming creative processes across industries. By presenting practical scenarios that prioritize interior design styles and emphasizing the synergy between AI, design knowledge, and visualization, this study contributes to the evolution of architectural and design practices. The contributions of the proposed methodology to the spatial design visualization approach are as follows:

(1) Enhanced spatial design visualization: AI models can create vivid and realistic visualizations from textual descriptions, aiding experts, nonexperts, and stakeholders in better understanding and communicating design concepts.
(2) Concept communication: Txt2Img AI allows experts to easily convey their ideas to nontechnical audiences. Moreover, complex spatial design visualization concepts can be translated into accessible visuals, fostering improved collaboration and understanding.
(3) Accessibility: AI-generated visualizations democratize access to spatial design visualization. Even those without advanced design software skills can create high-quality visuals using simple textual input.

Although this research focused on fine-tuning models for interior design styles, the methodology can be extended to accommodate diverse requirements based on regional, cultural, and user preferences. Furthermore, although the presented application is a demo, the future goal is to develop accessible applications that offer real and experiential usage, enhancing the potential of AI-supported visualization in architectural design.

Statement of conflicting interests

The authors state that they do not have any known financial interests or personal relationships that could have influenced the findings of this study.

Data availability statements

The first or corresponding author (PI) can provide most of the data for training and/or models that were used in this study upon a reasonable request, as well as the links in references and the technical resource section. In addition, the paper includes references to the archives and links provided by the PI.

Credit authorship contribution statement

Jin-Kook Lee: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing. **Hyun Jeong:** Formal analysis, Methodology, Validation, Writing – original draft, Writing – review & editing. **Youngchae Kim:** Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Seung Hyun Cha:** Formal analysis, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported in 2023 by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2021-KA163269). This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MIST) (No. NRF-2022R1A2C1093310).

References

[1] J. Abualdenien, A. Borrmann, Vagueness visualization in building models across different design stages, Adv. Eng. Inf. 45 (2020) 101107.
[2] Architectural Digest (2023). Retrieved from http://architecturaldigest.com.
[3] Arjovsky, M., Chintala, S., & Bottou, L. (2017, July) Wasserstein generative adversarial networks. In International conference on machine learning (pp. 214-223). PMLR.
[4] C. Ah-Soon, K. Tombre, Variations on the analysis of architectural drawings, Proce. Fourth Int. Conference on Document Analysis and Recognition 1 (1997) 347–351.
[5] O. Atilola, M. Tomko, J.S. Linsey, The effects of representation on idea generation and design fixation: a study comparing sketches and function trees, Des. Stud. 42 (2016) 110–136.
[6] A. Baldrati, D. Morelli, G. Cartella, M. Cornia, M. Bertini, R. Cucchiara, Multimodal garment designer: human-centric latent diffusion models for fashion image editing, arXiv preprint arXiv:2304.02051 (2023).
[7] Baty, P. (2017). The anatomy of colour: The story of heritage paints and pigments. Thames & Hudson.
[8] S. Bell, K. Bala, Learning visual similarity for product design with convolutional neural networks, ACM Trans. Graph. 34 (2015) 1–10.
[9] K. Borg, V. Sahadevan, V. Singh, T. Kotnik, Leveraging generative design for industrial layout planning: SWOT analysis insights from a practical case of papermill layout design, Adv. Eng. Inf. 60 (2024) 102375.
[10] H. Chang, H. Zhang, J. Barber, A.J. Maschinot, J. Lezama, L. Jiang, D. Krishnan, Muse: text-to-image generation via masked generative transformers, arXiv preprint arXiv:2301.00704 (2023).
[11] C.-S. Chan, Operational definitions of style, Environ. Planning B 21 (1994) 223–246.
[12] C.-S. Chan, Can style be measured? Des. Stud. 21 (2000) 277–291.
[13] F.D.K. Ching, A visual dictionary of architecture, John Wiley & Sons, 2011.
[14] M.L. Chiu, Collaborative design in CAAD studios: Shared ideas, resources, and representations, Proce. Int. Conference on CAAD Future 95 (1995) 749–759.
[15] J. Cherneff, R. Logcher, J. Connor, N. Patrikalakis, Knowledge-based interpretation of architectural drawings, Res. Eng. Des. 3 (1992) 195–210.
[16] C. Eckert, M. Stacey, Sources of inspiration: a language of design, Des. Stud. 21 (2000) 523–538.
[17] D. Epstein, A. Jabri, B. Poole, A.A. Efros, A. Holynski, Diffusion self-guidance for controllable image generation, arXiv preprint arXiv:2306.00986 (2023).
[18] R. Gal, Y. Alaluf, Y. Atzmon, O. Patashnik, A.H. Bermano, G. Chechik, D. Cohen-Or, An image is worth one word: personalizing text-to-image generation using textual inversion, arXiv preprint arXiv:2208.01618 (2022).
[19] E. Gates, Elements of style: Designing a home & a life, Simon & Schuster, 2014.
[20] L.A. Gatys, A.S. Ecker, M. Bethge, Image style transfer using convolutional neural networks, Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (2016) 2414–2423.
[21] G. Goldschmidt, Creative architectural design: reference versus precedence, J. Architectural and Planning Res. (1998) 258–270.
[22] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, Y. Bengio, Generative adversarial nets, Adv. Neural Inf. Proces. Syst. 27 (2014).
[23] Houzz (2009). Retrieved from http://www.houzz.com.
[24] E.J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, W. Chen, Lora: low-rank adaptation of large language models, arXiv preprint arXiv:2106.09685 (2021).
[25] Z. Hu, Y. Wen, L. Liu, J. Jiang, R. Hong, M. Wang, S. Yan, Visual classification of furniture styles, ACM Trans. Intelligent Systems and Technol. (TIST) 8 (2017) 1–20.
[26] K. Ikeno, T. Fukuda, N. Yabuki, An enhanced 3D model and generative adversarial network for automated generation of horizontal building mask images and cloudless aerial photographs, Adv. Eng. Inf. 50 (2021) 101380.
[27] Interior Design (2023). Retrieved from http://interiordesign.net.
[28] P. Isola, J.Y. Zhu, T. Zhou, A.A. Efros, Image-to-image translation with conditional adversarial networks, Proce. IEEE Conference on Comp. Vision and Pattern Recognition (2017) 5967–5976.
[29] Z. Jiang, Y. Ma, Y. Xiong, Bio-inspired generative design for engineering products: a case study for flapping wing shape exploration, Adv. Eng. Inf. 58 (2023) 102240.
[30] B. Jonson, Design ideation: the conceptual sketch in the digital age, Des. Stud. 26 (2005) 613–624.
[31] Johnson, J., Alahi, A., Fei-Fei, L. (2016) Perceptual losses for real-time style transfer and super-resolution. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14 (pp. 694-711). Springer International Publishing.
[32] D. Kalajdzievski, A rank stabilization scaling factor for fine-tuning with LoRA, arXiv preprint arXiv:2312.03732 (2023).
[33] T. Karras, S. Laine, T. Aila, A style-based generator architecture for generative adversarial networks, Proce. IEEE/CVF Conference on Comp. Vision and Pattern Recognition (2019) 4401–4410.
[34] B. Kawar, S. Zada, O. Lang, O. Tov, H. Chang, T. Dekel, I. Mosseri, M. Irani, Imagic: Text-based real image editing with diffusion models, Proce. IEEE/CVF Cnference on Comp. Vision and Pattern Recognition (2023) 6007–6017.
[35] Kilmer, R., Kilmer, W. O. (2014). Designing interiors. John Wiley & Sons: Hoboken, NJ, USA, 2014 (pp. 17-24).
[36] J. Kim, J.K. Lee, Stochastic detection of interior design styles using a deep-learning model for reference images, Appl. Sci. 10 (2020) 7299.
[37] J. Kim, J.-K. Lee, Auto-recognition of interior design images for managing architectural design references—Focused on the module implementation for recognizing the usage of rooms of Korean apartments, J. Korean Inst. Inter. Des 27 (2018) 13–20.
[38] Kim, J., Song, J., Lee, J. K. (2019). 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, (785-794).
[39] Kim, K., Park, S., Lee, J., Choo, J. (2023). Reference-based image composition with sketch via structure-aware diffusion model. arXiv preprint arXiv:2304.09748.
[40] Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012) Imagenet classification with deep convolutional neural networks. In Proceedings of the advances in neural information processing systems, lake Tahoe, NV, USA, 3–6 December 2012, pp. 1097-1105.
[41] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, W. Shi, Photo-realistic single image super-resolution using a generative adversarial network, Proc. IEEEConference on Comp. Vision and Pattern Recognition (2017) 4681–4690.
[42] J. Lee, J. Heo, J. Lee, Enhancement of virtual data quality using pre-trained Bayesian transfer learning under inaccurate and insufficient measurement data, Adv. Eng. Inf. 59 (2024) 102241.
[43] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444.
[44] M. Lin, T. Chen, H. Chen, B. Ren, M. Zhang, When architecture meets AI: a deep reinforcement learning approach for system of systems design, Adv. Eng. Inf. 56 (2023) 101965.
[45] X. Liu, C. Andris, Z. Huang, S. Rahimi, Inside 50,000 living rooms: an assessment of global residential ornamentation using transfer learning, EPJ Data Sci. 8 (2019) 4.
[46] P.E.D. Love, W. Fang, J. Matthews, S. Porter, H. Luo, L. Ding, Explainable artificial intelligence (XAI): precepts, models, and opportunities for research in construction, Adv. Eng. Inf. 57 (2023) 102024.
[47] A. Nichol, P. Dhariwal, A. Ramesh, P. Shyam, P. Mishkin, B. McGrew, M. Chen, Glide: towards photorealistic image generation and editing with text-guided diffusion models, arXiv preprint arXiv:2112.10741 (2021).
[48] Midjourney (2023). Midjourney (V5) [Text-to-image model]. https://www.midjourney.com/.
[49] P. Milgram, F. Kishino, A taxonomy of mixed reality visual displays, IEICE Trans. Inf. Syst. 77 (1994) 1321–1329.
[50] J. Oppenlaender, The creativity of text-to-image generation, Proce. 25th Int. Academic Mindtrek Conference (2022) 192–202.
[51] R. Oxman, Theory and design in the first digital age, Des. Stud. 27 (2006) 229–265.
[52] Pile, J. F. (1988). Interior Design.
[53] D.M. Phare, N. Gu, M. Ostwald, Representation in design communication: meaning-making in a collective context, Frontiers in Built Environ. 4 (2018) 36.
[54] J. Qu, Y. Zhang, W. Tang, W. Cheng, Y. Zhang, L. Bu, Developing a virtual reality healthcare product based on data-driven concepts: a case study, Adv. Eng. Inf. 57 (2023) 102118.
[55] A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, arXiv preprint arXiv:1511.06434 (2015).
[56] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, I. Sutskever, Zero-shot text-to-image generation, International Conference on Machine Learning (2021) 8821–8831 (PMLR).
[57] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, M. Chen, Hierarchical text-conditional image generation with clip latents, arXiv preprint arXiv:2204.06125 (2022).
[58] K.F. Richter, B. Weber, B. Bojduj, S. Bertel, Supporting the designer's and the user's perspectives in computer-aided architectural design, Adv. Eng. Inf. 24 (2010) 180–187.
[59] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-resolution image synthesis with latent diffusion models, Proce. IEEE/CVF Conference on Comp. Vision and Pattern Recognition (2022) 10674–10685.
[60] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, K. Aberman, Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation, Proce. IEEE/CVF Conference on Comp. Vision and Pattern Recognition (2023) 22500–22510.
[61] S. Sabeti, O. Shoghli, M. Baharani, H. Tabkhi, Toward AI-enabled augmented reality to enhance the safety of highway work zones: feasibility, requirements, and challenges, Adv. Eng. Inf. 50 (2021) 101429.
[62] R. Sacks, Z. Wang, B. Ouyang, D. Utkucu, S. Chen, Toward artificially intelligent cloud-based building information modelling for collaborative multidisciplinary design, Adv. Eng. Inf. 53 (2022) 101711.
[63] C. Saharia, W. Chan, S. Saxena, L. Li, J. Whang, E.L. Denton, M. Norouzi, Photorealistic text-to-image diffusion models with deep language understanding, Adv. Neural Inf. Proces. Syst. 35 (2022) 36479–36494.
[64] A.B. Saka, L.O. Oyedele, L.A. Akanbi, S.A. Ganiyu, D.W.M. Chan, S.A. Bello, Conversational artificial intelligence in the AEC industry: a review of present status, challenges and opportunities, Adv. Eng. Inf. 55 (2023) 101869.
[65] P. Sanguinetti, S. Abdelmohsen, J. Lee, J. Lee, H. Sheward, C. Eastman, General system architecture for BIM: An integrated approach for design and analysis, Adv. Eng. Inf. 26 (2012) 317–333.
[66] H.A. Simon, Style in design, in: C.M. Eastman (Ed.), Spatial Synthesis in Computer-Aided Building Design, Wiley:, New York, NY, USA, 1975, pp. 287–309.
[67] K. Sohn, N. Ruiz, K. Lee, D.C. Chin, I. Blok, H. Chang, D. Krishnan, StyleDrop: text-to-image generation in any style, arXiv preprint arXiv:2306.00983 (2023).
[68] A. Speltz, Styles of ornament: A collection of 4,000 architectural ornaments, Gramercy Books, 1994.
[69] Sutherland, I. E. (1964) Sketch pad a man-machine graphical communication system. In Proceedings of the share design automation workshop (pp. 6.329-6.346).
[70] Taigman, Y., Yang, M., Ranzato, M. A., Wolf, L. (2014) Deep face: Closing the gap to human-level performance in facenverification. In Proceedings of the IEEE

conference on computer vision and pattern recognition, Columbus, OH, USA (24-27 June 2014) (pp. 1701-1708).
[71] M. Uhm, G. Lee, Y. Park, S. Kim, J. Jung, J.K. Lee, Requirements for computational rule checking of requests for proposals (RFPs) for building designs in South Korea, Adv. Eng. Inf. 29 (2015) 602–615.
[72] Z. Wang, P. Zheng, X. Li, C.H. Chen, Implications of data-driven product design: from information age towards intelligence age, Adv. Eng. Inf. 54 (2022) 101793.
[73] K. Xu, X. Kong, Q. Wang, S. Yang, N. Huang, J. Wang, A bearing fault diagnosis method without fault data in new working condition combined dynamic model with deep learning, Adv. Eng. Inf. 54 (2022) 101795.
[74] Z. Ye, D. Li, J. Tian, T. Lan, J. Zuo, L. Duan, M. Tang, ASPEN: high-throughput LoRA fine-tuning of large language models with a single GPU, arXiv preprint arXiv:2312.02515 (2023).
[75] J. Yu, Y. Xu, J.Y. Koh, T. Luong, G. Baid, Z. Wang, Y. Wu, Scaling autoregressive models for content-rich text-to-image generation, arXiv preprint arXiv:2206.10789 2 (2022) 5.
[76] Yigitbas, E., Nowosad, A., & Engels, G. (2023). Supporting construction and architectural visualization through BIM and AR/VR: A systematic literature review. arXiv preprint arXiv:2306.12274.
[77] J.Y. Zhu, T. Park, P. Isola, A.A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, Proce. IEEE Int. Conference on Comp. Vision (2017) 2242–2251.