IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# Room Decor Ai Generated Interior Designing Using Stable Diffusion

Anagha Sajeev[1], Anand B[2], Eleena Mareyam George[3] and Dr. Rajesh K S[4]

Department Of Computer Applications,

Saintgits College Of Engineering, Pathamuttom, Kottayam

**Abstract:** This paper presents an AI-driven interior design system utilizing the Stable Diffusion model, developed by Stability AI and fine-tuned for design applications. The system generates personalized design recommendations by integrating user preferences and room photographs, producing layouts that enhance each room's structure. By automating much of the design process, it reduces the time and resources required by human expertise. The model was fine-tuned on a diverse dataset covering various room types, including bedrooms, living rooms, dining rooms, and kitchens, and supports styles like modern, minimalist, and rustic. Achieving an average accuracy of seventy-five in aligning with user input, the technology enhances creativity and workflow efficiency in design. Beyond interior spaces, it holds potential for broader design applications, providing a valuable tool for designers and homeowners to visualize room transformations. The system successfully generated design layouts that closely matched user preferences, demonstrating its effectiveness in creating personalized and aesthetically pleasing room transformations, with high user satisfaction in the final results.

**Index Terms:** interior design, stable diffusion, text generation design, artificial intelligence

## I. INTRODUCTION

There is a huge demand for interior design worldwide, but existing design approaches or methodologies may not fully meet these needs. One reason for this phenomenon is that the interior design process is complicated, and frequent changes lead to low design efficiency. In addition, designers form fixed design methods to save time, resulting in a lack of innovation. Hence, it is important to improve the efficiency of interior design and address the lack of innovation. With the introduction of the diffusion model, it is possible to solve the problems of low efficiency and a lack of creativity in interior design. The advantage of the diffusion model is that it can learn prior knowledge from the massive image and text description pairing information. The trained diffusion model can generate high-quality and diverse images by inputting text descriptions [2]. The interior design system introduced in this paper utilizes a fine-tuned stable diffusion model. Users can upload images of their rooms and choose how they want them by specifying design preferences through prompts. The system then generates realistic, customized room layouts that align with the structural features of the space, providing an alternative to traditional design approaches. Although AI has significantly influenced design, existing diffusion models encounter difficulties when applied to interior design, particularly in terms of achieving pixel-level alignment. This refers to ensuring that the generated elements such as furniture, decorations, or layout adjustments accurately match the room's structural characteristics in the original image. Without proper alignment, the generated design might appear unrealistic or impractical. This paper aims to address these limitations by presenting a fine-tuned stable diffusion model that generates creative yet practical designs in real-time. By ensuring that the design output aligns with the room's structure, the system reduces manual effort and enables both designers and non-designers to rapidly generate and modify designs

The specific objectives of this study are as follows:

1. To create an AI-powered system that uses user input and uploaded room photos to create personalized interior designs.
2. To adjust the Stable Diffusion model to fit different design styles, including modern, minimalist, and rustic, and to match room structures.
3. To assess how well the system matches user-specified design prompts in terms of precision, originality, and efficiency.
4. To contribute to the larger field of intelligent design automation by investigating the system's generalizability to other design-related tasks.

The advancements in generative AI, particularly diffusion models, have opened up new possibilities in creative fields. Our study fills in a gap in AI-driven interior design by providing a quick and easy-to-use solution that creates high-quality designs for both professionals and regular people.

## II. LITERATURE REVIEW

The field of interior design has faced persistent challenges related to inefficiency, creativity, and the ability to meet diverse client needs. Recent advancements in artificial intelligence (AI), particularly diffusion models, have emerged as innovative solutions to these challenges, enabling designers to leverage technology for creative and efficient design generation. This literature review examines key studies that demonstrate the transformative potential of AI diffusion models in the realm of interior design. The field of interior design is increasingly confronted with challenges related to inefficiency and aesthetic appeal, prompting the exploration of innovative technological solutions. Recent advancements in artificial intelligence (AI), particularly diffusion models, have emerged as transformative tools for generating creative designs from textual descriptions. [1] present a method for room interior generation using Stable Diffusion models, demonstrating the ability to quickly adapt to recent concepts without extensive retraining. This foundational work highlights the efficiency gains achievable through AI in the creative design process. In a complementary study, [2] integrate aesthetics and efficiency in AI-driven diffusion models, allowing for rapid generation of visually appealing interior designs based on specified text descriptions. Their findings underscore the importance of incorporating decorative styles and spatial functions, enhancing both design quality and efficiency. Furthering this exploration, [3] investigate text-to-image generation through a Stable Diffusion model combined with a variational autoencoder decoder, significantly improving the realism and fidelity of generated images. Expanding on image manipulation, [4] introduce a plug-and-play diffusion feature framework for image-to-image translation, empowering users with greater control over generated content. This flexibility is vital in the context of interior design, where precise adjustments are often necessary.[5] also propose a novel diffusion model-based method for generating interior designs from text, emphasizing the potential for efficiency in creative design generation, enabling quick adaptations to the client needs. In a broader context, [6] refine text-to-image generation techniques to enhance accuracy in AI-generated designs. Their work illustrates how conditional controls can improve the alignment of generated images with user specifications.[7] demonstrate the effectiveness of pixel-aware diffusion for realistic image super-resolution and personalized stylization, crucial for tailoring interior design solutions to specific client preferences. Exploring the intersection of control and generative capabilities,[8] introduce Uni-ControlNet, an all-in-one framework that facilitates advanced manipulation of text-to-image diffusion models. This development provides designers with unprecedented control over the generated content, further streamlining the design process.[9] present DreamPose, a method that synthesizes fashion images into video, indicating the potential for dynamic and multi-dimensional applications of AI in design. In addition, [10] discuss MultiDiffusion, which fuses diffusion paths for controlled image generation, enhancing the creative possibilities available to designers. The importance of integrating conditional control into text-to-image diffusion models is further emphasized by [11], who highlights the potential for fine-tuning in achieving desired aesthetic outcomes. Research by [12] explores the concept of chain-of-thought reasoning in AI systems, suggesting that such reasoning can enhance the decision-making capabilities of design models.[13] investigate the role of ensemble expert denoisers in improving text-to-image diffusion models, reinforcing the significance of advanced technical methodologies in achieving high-quality design outputs. Additionally,[14] focus on DreamBooth, a technique for fine-tuning text-to-image diffusion models for subject-driven generation, providing personalized design outputs that cater to specific client needs.[15] further contribute to this discourse by discussing high-resolution image synthesis with latent diffusion models, which underscores the potential for producing intricate and detailed designs. The exploration of visual and vision-language representation learning is advanced by [16], highlighting the role of noisy text supervision in enhancing generative models.[17] lay foundational work in denoising diffusion probabilistic models, providing critical insights into the underlying mechanisms that drive these technologies .Lastly, [18] examine the application of image style transfer technology in interior decoration design, illustrating the potential for blending traditional design techniques with modern AI capabilities. [19] conclude this extensive discourse by proposing a diffusion model with an improved control network, which facilitates the generation of creative interior designs that align with indoor structures. The integration of AI diffusion models in interior design enhances creativity, efficiency, and personalization, enabling designers to generate visually appealing and tailored spaces. This approach offers significant improvements in adapting to client needs and transforming traditional design processes.

## III. METHODOLOGY

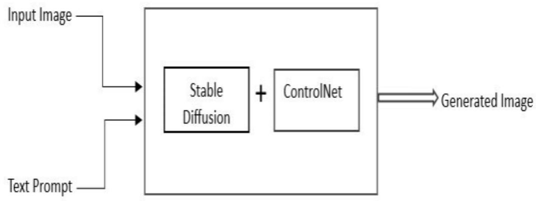
The methodology of the proposed system is organized into four primary modules: Data Collection and Preprocessing Module, ControlNet and Stable Diffusion Pipeline Module, Fine-Tuning Module and Inference Module.

### A. Data Collection and Preprocessing Module

The initial module loads room images from a specified folder, pairing them with corresponding text prompts stored in `.txt` files. Each text file provides a description of the room, creating an organized image-text pair dataset. This setup enables users to define design preferences through descriptive prompts, forming the foundation for the system's AI-driven design generation. Images are pre-processed via resizing, normalization, and transformation into tensors, ensuring a standardized input format for model training and inference.

B. ControlNet and Stable Diffusion Pipeline Module Serving as the system's core, this module combines structure-preserving techniques with advanced generative models to produce customized room designs. ControlNet, a neural network designed for adding constraints to diffusion models, leverages Canny edge-detected images to enhance spatial accuracy, thereby preserving room layout and architectural features. Stable Diffusion, fine-tuned with DreamBooth techniques on Moroccan interiors (referred to as "beldi-v2," trained by medmac01 on the "medmac01/dreambooth-moroccan-design-v2" dataset), manages the creative aspects of design generation. Together, ControlNet ensures layout fidelity, while Stable Diffusion produces visually cohesive, culturally specific designs based on user prompts. This combined approach delivers personalized, structurally accurate, and visually cohesive interior designs tailored to user preferences.

*Figure 1 System architecture*




C. Fine-Tuning Module

During fine-tuning, each image is paired with its corresponding text file and processed through the pipeline to generate an initial design. The Mean Squared Error (MSE) loss function calculates the difference between the generated and input images, with the system aiming to minimize this loss over training epochs to enhance design quality. Any black images detected, often indicating undesirable outputs, are skipped to avoid negative effects on training. The Adam optimizer refines model weights iteratively, reducing the average loss per epoch. The module logs loss per batch, calculates the average loss per epoch, and generates a loss trend plot for easy assessment. Upon completion of the specified epochs, the fine-tuned model is saved, preserving optimized parameters for future inference.

D. Inference Module

Designed for real-time, user-friendly design generation, the Inference Module allows users to quickly create custom room designs. Users upload a room image and specify a design prompt, providing tailored input for the model. The uploaded image undergoes preprocessing, where it is resized to 256x256 pixels and subjected to Canny edge detection to extract structural outlines that maintain the room's layout. This processed image, combined with the user's prompt, is then passed through the fine-tuned ControlNet and Stable Diffusion pipeline. ControlNet preserves the room layout, while Stable Diffusion generates a design aligned with the prompt, delivering a customized interior design that respects both room structure and user preferences. The generated design is displayed in real time for user review and saved for future reference.

Data collection and preprocessing standardize room images and prompts. ControlNet uses Canny edge detection to maintain structural alignment, while Stable Diffusion generates designs based on the prompts. The MSE loss function measures output quality, guiding model refinement. The Fine-Tuning Module orchestrates training, logging loss per epoch, saving the fine-tuned model, and monitoring progress through loss visualization. The Inference Module manages image uploads, preprocessing, and design generation, ensuring layout-aligned outputs and filtering unwanted content. These interconnected modules together deliver an efficient, AI-driven pipeline for generating interior designs that balance ControlNet and Stable Diffusion for high-quality, structure-aligned outputs tailored to user inputs.

The images generated in the training loop are the result of processing input images through the **Stable Diffusion ControlNet Pipeline**. This pipeline integrates two key components: **Stable Diffusion** for image generation and **ControlNet** for adding structural guidance. Each of these components plays an essential role in ensuring that the generated images are both visually coherent and aligned with the design prompts provided by the user. The dataset consists of pairs of input images and corresponding text prompts. Each image is linked with a prompt that describes the desired features and style of the output image. Further explained in Results and Discussion Section. The Custom Dataset is used to load the input images and their respective prompts. The images are resized to 512x512 pixels and converted into tensors, while the text prompts are loaded as strings. These inputs are then processed throughout the model's training loop. The image generation process involves the use of two models: **Stable Diffusion** and **ControlNet**. The **Stable Diffusion Pipeline** generates images based on text prompts. This model uses a diffusion process in which random noise is iteratively refined to produce meaningful images. The model gradually turns this noise into an image that aligns with the content described in the prompt. On the other hand, **ControlNet** adds an additional layer of control by incorporating external structural information, such as edges, poses, or depth maps, to guide the diffusion process. In this work, **Canny edge detection** is used to extract structural features from the input image, which are then utilized by ControlNet to guide the generation process. This ensures that the generated images retain the structural integrity of the input while adapting to the content and style described by the prompt. During each iteration of the training loop, the following sequence of steps occurs. First, an image from the dataset is processed by converting it into a tensor and transferring it to the GPU for processing. The image tensor is then converted into PIL image format, as required by the Stable Diffusion pipeline. The text prompt, along with the processed image, is passed into the StableDiffusionControlNetPipeline for image generation. This model takes into account both the structural features derived from the input image and the stylistic elements defined in the prompt. The model uses **Canny edge detection** to retain the essential edges and structure of the original image, while **Stable Diffusion** refines the output to reflect the desired style. A **negative prompt** is also provided to avoid generating undesirable features, and the process is iterated for a set number of steps (e.g., 20 steps), refining the image during each step. Once the image is generated, it is saved to disk with filenames reflecting the epoch and batch, allowing for easy tracking of the model's progress throughout training. To assess the quality of the generated images, a **Mean Squared Error (MSE)** loss function is used. This loss is computed by comparing the generated image with the original input image from the dataset. The MSE quantifies the difference between the generated image and the input image, guiding the model toward minimizing this difference over the course of training. This loss function is critical in refining the model's ability to generate images that closely align with both the structural and stylistic aspects described in the prompt. The images produced during training reflect the combined influence of the input image and the text prompt. ControlNet ensures that the structural content of the generated images aligns with the input image by leveraging edge information, while Stable Diffusion adapts the content to match the stylistic elements specified in the prompt. The output images demonstrate the model's ability to blend both the content and style effectively.

In conclusion, the image generation process begins with loading the input image and its corresponding prompt. ControlNet uses edge detection to guide the generation process, ensuring that the structural elements of the input are preserved. Stable Diffusion then generates a new image based on the text prompt, refining it over multiple inference steps. The generated image is evaluated using a loss function that compares it to the original input image. After several epochs, the model continues to refine the output, resulting in images that successfully blend the content of the input with the stylistic influence of the prompt. This process highlights the potential of combining text-based guidance with structural information to produce high-quality, contextually relevant images.

## IV. RESULT AND DISCUSSION

### A. Datasets

The performance of the proposed system is evaluated using a custom dataset we specifically collected for this system. The dataset consists of image-text pairs, enabling the system to be fine-tuned for generating high-quality images based on detailed textual prompts. The dataset is designed to cover various scenarios, ensuring diversity in input data for robust training and testing. These image-text pairs provide essential data for training the fine-tuned Stable Diffusion model with ControlNet. The dataset includes images captured from various angles, lighting conditions, and object compositions, ensuring a diverse representation of real-world scenes. Each image is carefully annotated with descriptive prompts that convey the visual content and context, making the dataset ideal for training a text-to-image generation model. The image resolutions are 256x256 pixels, balancing quality and computational efficiency for training and inference. This proprietary dataset plays a crucial role in ensuring that the proposed system accurately learns the relationship between visual elements and their corresponding textual descriptions. By leveraging this dataset, the system can generate customized images that align closely with user prompts, demonstrating versatility in a variety of image generation scenarios. The collection and annotation process of this dataset was carried out to ensure high-quality, detailed data, allowing for effective fine-tuning of the model. Figure 1 and Table 1 show the image-text pair custom dataset collected for this system, where Image 1 in Figure 1 corresponds to Text 1 in Table 1, and so on.

Figure 3 Sample of custom image datasets

Table 1 Sample of custom text datasets

| Room | Text |
|---|---|
| Bedroom 1 | A modern minimalist bedroom with a large white bed, simple black furniture, and abstract art on the wall, creating a clean and contemporary look. |
| Bedroom 2 | A colorful and playful children's bedroom with bright walls, a multicolored bedspread, and toy organizers, giving the space a fun and lively atmosphere. |
| Bedroom 3 | A simple twin-bed guest room with blue bedding, soft lighting, and minimal decor, providing a cozy and functional space for visitors. |
| Bedroom 4 | A spacious master bedroom with large windows, a queen-sized bed, and pastel blue walls, creating a calm and serene environment. |

| | |
|---|---|
| Kitchen 1 | A modern kitchen with sleek wooden cabinets, a corner countertop, and stainless steel appliances, combining style with functionality. |
| Kitchen 2 | A rustic and traditional kitchen featuring worn wooden countertops, vintage storage shelves, and small kitchen tools, giving it an old-world charm. |
| Kitchen 3 | A contemporary kitchen with white cabinetry, a granite island, and stainless steel appliances, showcasing a clean and modern design. |
| Kitchen 4 | A spacious and bright kitchen with white cabinets, a wooden countertop, and large windows providing ample natural light, creating an airy and fresh atmosphere. |
| Dining room 1 | A formal dining room with elegant wooden furniture, a chandelier, and large windows that bring in plenty of sunlight, perfect for family gatherings. |
| Dining room 2 | A cozy bar area with high stools, warm lighting, and rustic wooden cabinets, offering a welcoming spot for socializing. |
| Dining room 3 | A large open kitchen with a long dining table, wooden cabinets, and high ceilings, combining rustic and modern elements. |
| Dining room 4 | A bright dining space with a simple wooden dining set, contemporary art on the walls, and warm ambient lighting for a modern and welcoming vibe. |
| Living Room 1 | A luxurious living room with a spacious sectional sofa, modern decor, and soft lighting, offering a sleek and comfortable setting. |
| Living Room 2 | A cozy living room with terracotta-tiled floors, wooden furniture, and a large map hanging on the wall, providing a homey and relaxed atmosphere. |
| Living Room 3 | A traditional living room with a large aquarium, wooden decor, and plush seating, creating a warm and inviting environment. |
| Living Room 4 | A minimalist living space with light gray walls, a beige sofa, and natural wood flooring, combining simplicity and comfort. |

B. Fine Tuning

The model was fine-tuned over different epochs. The Mean Squared Error (MSE) was employed as the loss function to measure the difference between the generated images and the input images. Images were resized to 256x256 pixels and normalized to facilitate consistent input. Upon completion of training, the model was evaluated using test inputs to assess its ability to generate images based on the design prompts**.** The generated images displayed improvements in clarity and detail compared to initial outputs. The integration of Canny Edge Detection helped preserve structural integrity, resulting in aesthetically pleasing images. The model's ability to interpret and execute prompts was assessed qualitatively. In various cases, the generated images closely aligned with the specified design prompts, showcasing the model's understanding of the input text.

Figure 4 Sample of canny edges detection of input images

While the model demonstrated promising results, several limitations were noted such as, Black Images Detection, Some generated images were flagged as "black" or contained unwanted artifacts. The implemented checks effectively filtered these outputs, ensuring higher quality results. Sensitivity to Prompts, The model occasionally struggled with ambiguous prompts, resulting in varied outputs that did not align with user expectations.

After fine-tuning, the model was saved and used to create an API with Flask. We tested this API using HTTPie to ensure its stability, then integrated it with our frontend application. This setup allows us to generate images dynamically based on user-selected design prompts, providing an interactive experience for users to explore various interior design styles. This streamlined API connection ensures the generated images are quickly available for real-time use in the frontend, enhancing responsiveness and user engagement.

## C. EVALUATING EFFECTIVENESS

In Example Set 1 and Example Set 2, the user inputs both images and their desired design style as prompts. For each set, the user provides a reference image and a description of the desired style, such as "a modern, minimalist dining room with sleek furniture and neutral tones" or "a cozy, rustic living room with warm lighting and wooden accents." The fine-tuned model evaluates the input images alongside the style prompts, incorporating Canny Edge Detection to preserve the key structural elements and edges within the design. This process helps ensure that the critical features of the design, such as furniture placement and room layout, are maintained while applying the desired stylistic elements. Based on this evaluation, the model generates an output image that reflects the user's input, showcasing the transformed design with improved detail, clarity, and stylistic coherence. These generated images are then made accessible through the Flask API, which is tested with HTTPie to ensure that the process works seamlessly, and the final image is displayed in real-time on the frontend.

1. Example Set 1

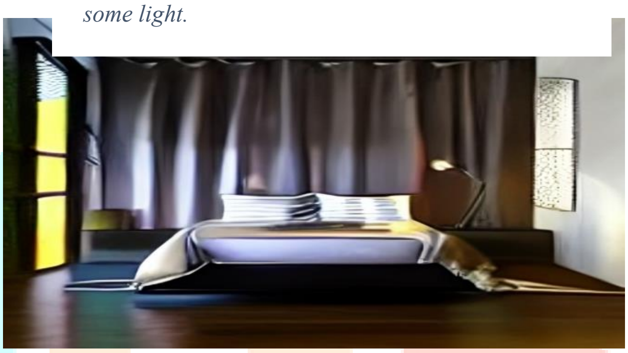
*Figure 5 Input image*

2. Example Set 2

*Figure 6 Image generated after the prompt - Redesign the room with minimalist design with light colours*

*Figure 7 Input image*

*Figure 8 Image generated after text prompt – Redesign the room by adding curtains change bed sheet and add some light.*



The analysis of the example sets demonstrates that the fine-tuned Stable Diffusion model effectively generates high-quality images closely aligned with the provided design prompts. The model showcased improved clarity and detail, successfully interpreting diverse themes, such as minimalist and vibrant children's rooms. However, some limitations were observed, including the occasional generation of black images and sensitivity to ambiguous prompts. Addressing these issues will enhance the model's reliability. Overall, the results indicate a promising foundation for customized image generation in interior design, with potential for further refinement and application.

## V. CONCLUSION

This paper presents an AI-driven approach for generating customized interior designs by integrating ControlNet and Stable Diffusion. The system uses a custom dataset of image-text pairs for high-quality image generation based on user prompts. It comprises four modules: Data Collection and Preprocessing, ControlNet and Stable Diffusion Pipeline, Fine-Tuning, and Inference Module. The first module organizes room images and prompts for model training, While the main pipeline integrates structure-preserving methods with generative models to guarantee spatial accuracy and originality. Fine-tuning, guided by the Mean Squared Error loss function, enhances model alignment with user prompts. The Inference Module allows real-time design generation. Although issues like dark image detection and sensitivity to unclear instructions still exist, the system effectively creates designs that are in line with user inputs while maintaining structural integrity through Canny Edge Detection. By combining creativity and structural accuracy, this research advances the automation of interior design and opens up possibilities for applications in architecture and design. Expanding the dataset to include a wider variety of scenarios and styles could further enhance the model's versatility and applicability across diverse interior design contexts. In conclusion, the proposed AI-driven system represents a significant step toward automating the interior design process, combining artistic creativity with structural fidelity. By providing users with the ability to create personalized designs through intuitive prompts, this research not only demonstrates the capabilities of modern generative models but also opens up new possibilities for future applications in interior design and architecture.

## VI. REFERNECES

[1] Shubh Sahu, Shreya Verma, Om Patil, Aryan Ghogare, and U. Kumar, "Room Interior Generation Using Stable Diffusion Models," International Journal For Multidisciplinary Research, vol. 6, no. 3, May 2024, doi: https://doi.org/10.36948/ijfmr.2024.v06i03.18627.

[2] J. Chen, Z. Shao, X. Zheng, K. Zhang, and J. Yin, "Integrating aesthetics and efficiency: AI-driven diffusion models for visually pleasing interior design generation," Scientific Reports, vol. 14, no. 1, p. 3496, Feb. 2024, doi: https://doi.org/10.1038/s41598-024-53318-3.

[3] J. Chen et al., "Creative interior design matching the indoor structure generated through diffusion model with an improved control network," Frontiers of Architectural Research, Sep. 2024, doi: https://doi.org/10.1016/j.foar.2024.08.003.

[4] S. Sahu, A. Ghogare, U. Sharma, S. Verma, and O. Patil, "Room Interior Generation Using Stable Diffusion Models," arXiv preprint, 2024.

[5] U. Budige and S. G. Konda, "Text-to-Image Generation Using Stable Diffusion Model with Variational Autoencoder Decoder," arXiv preprint, 2023.

[6] N. Tumanyan, M. Geyer, S. Bagon, and T. Dekel, "Plug-and-Play Diffusion Features for Text-Driven Image-to-Image Translation," arXiv preprint, 2022.

[7] J. Chen et al., "Creative interior design matching the indoor structure generated through diffusion model with an improved control network," Frontiers of Architectural Research, Sep. 2024, doi: https://doi.org/10.1016/j.foar.2024.08.003.

[8] Yang, T., Wu, R., Ren, P., Xie, X., & Zhang, L. (2024). Pixel-Aware Stable Diffusion for Realistic Image Super-Resolution and Personalized Stylization. European Conference on Computer Vision (ECCV) 2024.

[9] Zhao, S., Chen, D., Chen, Y., Bao, J., Yuan, L., Hao, S., & Wong, K. Y. K. (2023). Uni-ControlNet: All-in-One Control to Text-to-Image Diffusion Models. arXiv preprint.

[10] J. Chen et al., "Creative interior design matching the indoor structure generated through diffusion model with an improved control network," Frontiers of Architectural Research, Sep. 2024, doi: https://doi.org/10.1016/j.foar.2024.08.003.

[11] Bar-Tal, O., Yariv, L., Lipman, Y., & Dekel, T. (2023). MultiDiffusion: Fusing Diffusion Paths for Controlled Image Generation. Frontiers of Architectural Research.

[12] Zhang, L., Rao, A., & Agrawala, M. (2023). Adding Conditional Control to Text-to-Image Diffusion Models. arXiv preprint.

[13] Lyu, Q., Havaldar, S., Stein, A., Zhang, L., Rao, D., Wong, E., & Apidianaki, M. (2023). Faithful Chain-of-Thought Reasoning. arXiv preprint.

[14] Balaji, Y., Nah, S., Huang, X., Vahdat, A., Zhang, Q., Kreis, K., Aittala, M., Laine, S., Catanzaro, B., Karras, T., & Liu, M. Y. (2023). eDiff-I: Text-to-Image Diffusion Models with an Ensemble of Expert Denoisers. arXiv preprint.

[15] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, and K. Aberman, "DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation."

[16] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, and K. Aberman, "DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation"

[17] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. arXiv preprint.

[18] Jia, C., Yang, Y., Xia, Y., Chen, Y. T., Parekh, Z., Pham, H., Le, Q. V., Sung, Y., Li, Z., & Duerig, T. (2021). Scaling Up Visual and Vision-Language Representation Learning with Noisy Text Supervision. Frontiers of Architectural Research.

[19] Ho, J., Jain, A., & Abbeel, P. (2020). Denoising Diffusion Probabilistic Models. Frontiers of Architectural Research.