buildings

MDPI

Article

Generating Interior Design from Text: A New Diffusion Model-Based Method for Efficient Creative Design

Junming Chen [1,†], Zichun Shao [1,†] and Bin Hu [1,2,*]

[1] Faculty of Humanities and Arts, Macau University of Science and Technology, Macao 999078, China; jmchen@must.edu.mo (J.C.); 2220015871@student.must.edu.mo (Z.S.)
[2] School of Arts, Soochow University, Suzhou 215006, China
* Correspondence: binhu@must.edu.mo
† These authors contributed equally to this work.

Abstract: Because interior design is subject to inefficiency, more creativity is imperative. Due to the development of artificial intelligence diffusion models, the utilization of text descriptions for the generation of creative designs has become a novel method for solving the aforementioned problem. Herein, we build a unique interior decoration style dataset. Thus, we solve the problem pertaining to the need for datasets, propose a new loss function that considers the decoration style, and retrain the diffusion model using this dataset. The trained model learns interior design knowledge and can generate an interior design through text. The proposed method replaces the designer's drawing with computer-generated creative design, thereby enhancing the design efficiency and creative generation. Specifically, the proposed diffusion model can generate interior design images of specific decoration styles and spatial functions end to end from text descriptions, and the generated designs are easy to modify. This novel and creative design method can efficiently generate various interior designs, promote the generation of creative designs, and enhance the design and decision-making efficiency.

Keywords: diffusion model; text generation design; interior design; artificial intelligence; efficient; creative design

check for updates

Citation: Chen, J.; Shao, Z.; Hu, B. Generating Interior Design from Text: A New Diffusion Model-Based Method for Efficient Creative Design. *Buildings* **2023**, *13*, 1861. https://doi.org/10.3390/buildings13071861

Academic Editor: Svetlana J. Olbina

Received: 8 June 2023
Revised: 14 July 2023
Accepted: 20 July 2023
Published: 22 July 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

There is a huge demand for interior design worldwide, but existing design approaches or methodologies may not fully meet these needs [1–3]. One reason for this phenomenon is that the interior design process is complicated, and frequent changes lead to low design efficiency [1,3–5]. In addition, designers form fixed design methods to save time, resulting in a lack of innovation [1,6,7]. Therefore, it is important to improve the efficiency of interior design and address the lack of innovation.

With the introduction of the diffusion model [8,9], it is possible to solve the problems of low efficiency and a lack of creativity in interior design [10–12]. The advantage of the diffusion model is that it can learn prior knowledge from massive image and text description pairing information [13–16]. The trained diffusion model can generate high-quality and diverse images by inputting text descriptions in batches. Using the diffusion model for interior design can generate design schemes in batches for designers. This method can effectively improve the efficiency of design and creative generation [17–20].

Although diffusion models work well in most domains, they generate poor image quality in a few domains. Especially in the field of interior design, which requires highly professional skills, conventional diffusion models cannot produce high-quality interior designs. For example, the current mainstream diffusion models Midjourney [21], Dall E2 [13], and Stable Diffusion [22] cannot generate high-quality design images with specified decoration styles and space functions (Figure 1). The correct decoration style and space function are very important to interior design, and thus it is urgent to solve the above problems.



Figure 1. Images generated by mainstream diffusion models are compared with those generated by our method. Midjourney [21] will produce a lot of redundant objects, and the image is not realistic (**left one**). The object generated by DALL E2 is incomplete and has the wrong space size (**second from the left**). The placement and spatial scale of things generated by Stable Diffusion [22] are incorrect (**third from the left**). None of these images are up to the interior design requirements, and our proposed method (**far right**) improves the above problems. (prompt word: "Realistic, Chinese-style study room, with desks and cabinets").

In order to batch-generate designs with specific decoration styles and space functions, this study created a new interior decoration style dataset and retrained the diffusion model to make it suitable for interior design generation. Specifically, this study first collected a brand new Interior Decoration Style and Space Function (IDSSF-64) dataset from professional designers to solve the problem of a lack of training datasets for this task. IDSSF-64 includes the classification of decoration styles and space functions. Then, we proposed a new loss function, which adds style-aware reconstruction loss and style prior preservation loss to the conventional loss function. This function forces the diffusion model to learn the knowledge of decoration styles and space functions and retains the basic knowledge of the original model. The new model proposed in this study uses a new loss function and a new dataset to fine-tune training for interior design generation with specified decoration styles and space functions. The fine-tuning method does not need to retrain the whole model. It only requires a small number of images to fine-tune the model to obtain a better generation effect, thus significantly reducing the amount of training data and training time. The fine-tuned model can generate end-to-end interior designs in batches for designers to select, thereby improving design efficiency and creativity. The framework of this study is shown in Figure 2.

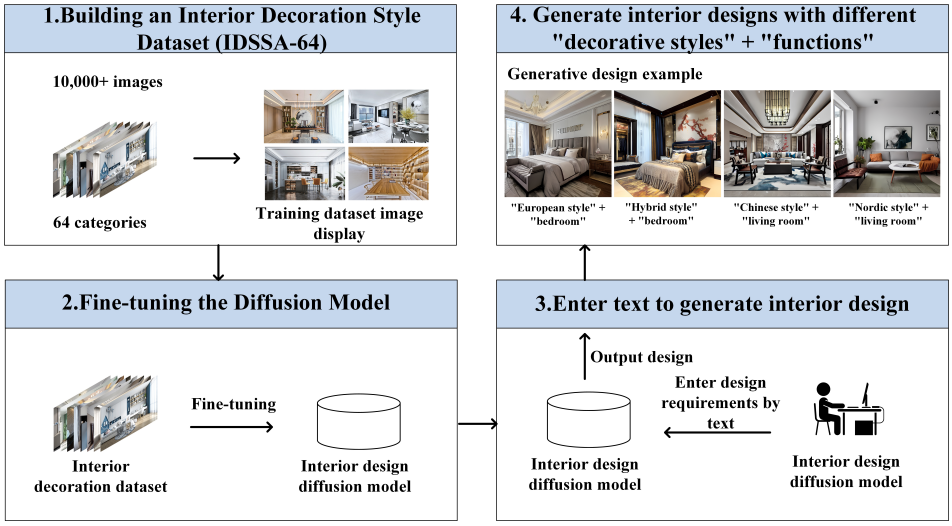


Figure 2. Study framework. This study first collects the interior decoration style dataset IDSSF-64 and then builds a diffusion model suitable for interior design through fine-tuning. Users can input a decoration style and space function into the fine-tuned diffusion model to directly obtain the design.

The fine-tuned diffusion model generative design method proposed in this study has changed the design process, and interior design efficiency and creativity have improved. The model can generate a variety of indoor spaces and ensure that the generated content meets the design requirements. Figure 3 demonstrates the interior design effects of different decoration styles and spaces generated by our model. The figure shows that the model understands the decoration styles and space functions. Each generated object appears in a suitable position, resulting in high-quality interior design.

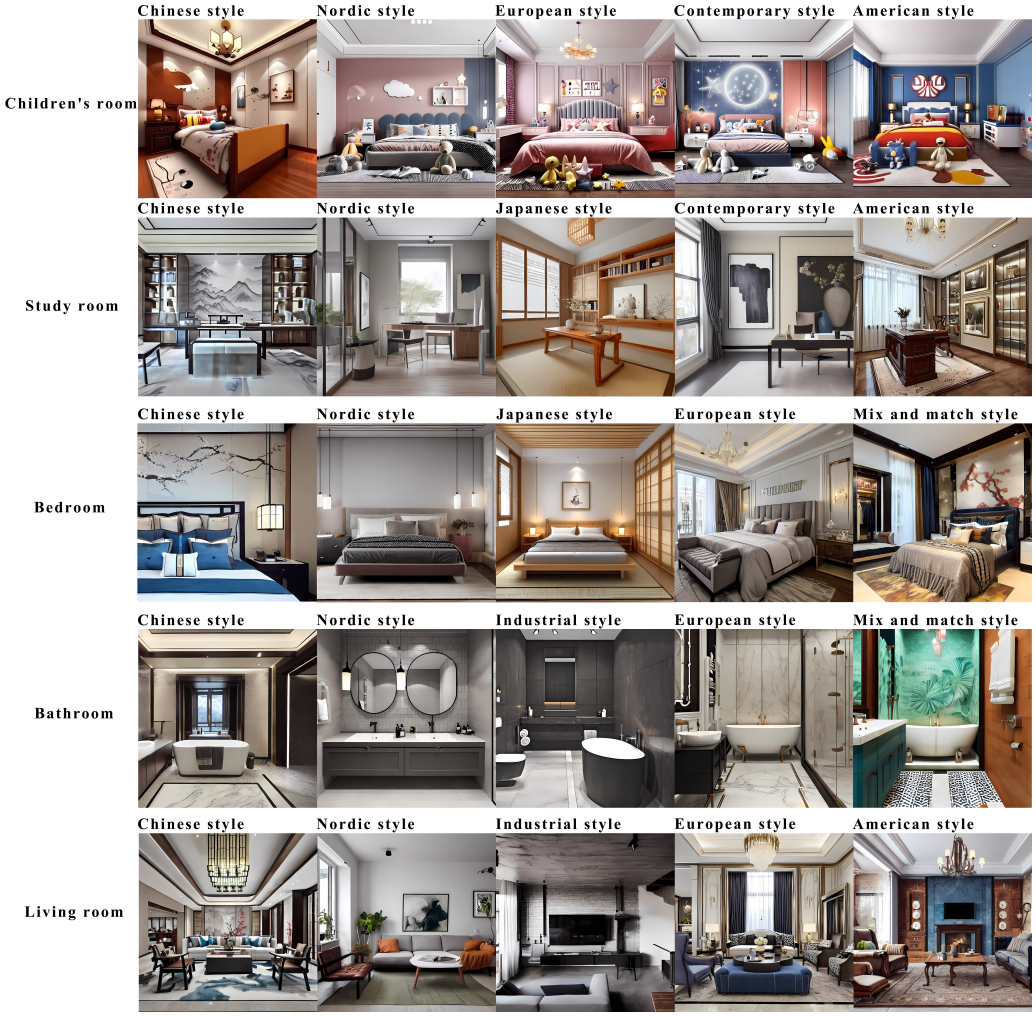


Figure 3. Interior design images generated by our diffusion model for different decoration styles and space functions.

The main contributions of this study are the following:

1. Proposing a novel end-to-end interior design method that directly generates designs from textual descriptions;
2. Proposing a new loss function and retraining the diffusion model to generate a diverse and high-quality interior design;
3. Creating a new indoor dataset with high precision, including decorative styles and space functions;
4. A comparison with other mainstream text-to-image diffusion models demonstrating the advantages of our method.

2. Literature Review

2.1. Conventional Interior Design Process

Interior design usually means that designers use their art and engineering knowledge to design interior spaces with specific decorative styles for clients. Designers must choose appropriate design elements to shape the decoration style, such as suitable tiles, furniture, colors, and patterns. A strong decoration style is key to making the design unique [3].

Designers usually use decorative renderings to determine the final design with clients, but this approach is inefficient. The reason for this inefficiency is that the conventional interior design workflow is linear, and designers spend a lot of time drawing design images and cannot communicate with customers in real time, resulting in many revisions. The conventional design process is shown in Figure 4. Specifically, interior design usually requires designers to find intentions to discuss with customers and decide on the decoration style. Then, the designer starts to produce two-dimensional (2D) drawings and build corresponding three-dimensional (3D) models. Then, material mapping is assigned to the 3D model, lighting is arranged for the space, and renderings are obtained. Finally, the customer determines whether the design is suitable by observing the renderings [5]. The linear workflow requires designers to design step by step. Once the customer is unsatisfied with a particular node, the designer must redo the entire design, leading to low design efficiency.

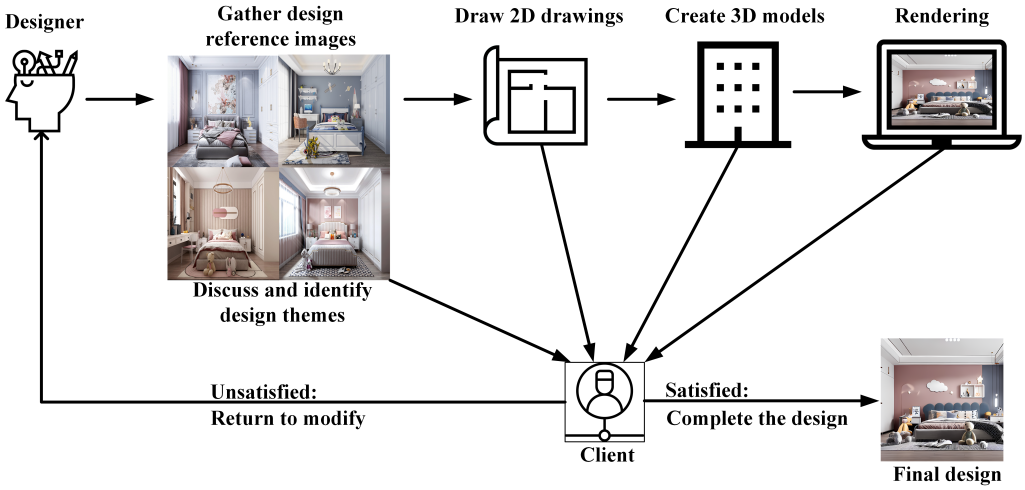


Figure 4. Conventional interior design process. Designers need to complete the design through a linear design process. If the customer is unsatisfied with the design during the process, then the designer must redo the entire design process.

At the same time, the cumbersome interior design workflow also suppresses creative design. On one hand, designers will form a fixed design method in the pursuit of efficiency so that they can quickly produce creative designs. On the other hand, even if the designer has a lot of creative inspirations, it takes a lot of labor to transform them into renderings, and they can only draw some of these within a limited time. Helping designers quickly obtain diversified interior design renderings is the key to solving the problems of low design efficiency and insufficient creativity.

Existing design automation mainly focuses on a particular process in the design, but fewer studies focus on end-to-end design [4,5]. This study achieves end-to-end generation of interior design by building a text-to-image diffusion model, thereby improving design efficiency and addressing the lack of creativity.

2.2. Text-to-Image Diffusion Model

The earlier diffusion model was proposed in 2015 [8] and has been continuously optimized and improved since then [23]. The improved model has become a new mainstream

generative model due to its excellent productive image effects [23–25]. The diffusion model mainly includes forward and reverse processes, and the forward process continuously adds noise to the original image. The reverse process iteratively denoises purely random noise to restore the image. Diffusion models learn the denoising process to gain the ability to generate images [8,23]. Generating an image of a specified category or with specific features requires adding text guidance. Text-to-image-based diffusion models enable controlled image generation using text as a guiding condition [10,11,25–27]. An advantage of the text-guided diffusion model is that it can create images that match the meanings of the prompt words.

There are two ways to learn new knowledge in the diffusion model: to retrain the entire model and to fine-tune the model to make the model suitable for new scenarios. Considering the high cost of retraining the whole model, fine-tuning the model is more feasible. There are four commonly used methods for fine-tuning models. The first one is textual inversion [13,25,28,29]. The core idea of textual inversion to embed new knowledge into a model is to freeze the text-to-image model and only give the most suitable embedding vector for new knowledge. This approach does not require model changes and is similar to finding new representations in the model to represent new keywords. The second is the hypernetwork [30]. A hypernetwork is a separate small neural network. The model is inserted into the middle layer of the original diffusion model to affect the output. The third is LoRA [31]. LoRA adds its weight to the attention cross layer as fine-tuning. The fourth is DreamBooth [32], which expands the text-image dictionary of the target model and establishes a new type of association between text identifiers and images while using rare words to name new knowledge and train to avoid language drift [33,34]. At the same time, we designed a prior preservation loss function to solve the overfitting problem. This loss function prompts the diffusion model to produce different examples of the same category as the subject. This method only needs 3–5 images and corresponding text descriptions to complete the fine-tuning of a specific theme and match the detailed text descriptions with the characteristics of the input image. The fine-tuned model can generate images with the trained topic words and descriptors. DreamBooth usually works best among these methods because it fine-tunes the entire model.

2.3. Public Dataset

Datasets are the basis for the rapid development of artificial intelligence, and many open datasets have promoted this. For example, the ImageNet dataset contains more than 10 million images, promoting the rapid development of computer vision [35]. Another example is the Common Objects in Context (COCO) dataset, which contains image recognition, segmentation, and semantic understanding data and has made significant progress in image segmentation and semantic understanding tasks in the field of computer vision [36]. At the same time, it has almost become a benchmark dataset for evaluating the performance of image semantic understanding algorithms. For example, the large-scale CelebFaces Attributes (CelebA) dataset, launched by the laboratory of Professor Tang Xiaoou of the Chinese University of Hong Kong, contains 200,000 face pictures and more than 40 kinds of face attributes, which promotes the development of face recognition tasks [37].

Text-to-image generative diffusion models have been slow to develop in interior design due to the lack of image-text paired interior datasets describing decorative styles and spatial functions. Specifically, the conventional diffusion model lacks suitable interior design datasets for training, and thus the diffusion model cannot stably generate designs for specified decoration styles and spatial functions. To this end, it is imminent to establish a new dataset of interior design decorating styles.

3. Methodology

Diffusion models based on image generation methods from the text have achieved impressive results in recent years, with CogView2 [38], Imagen [14], DreamBooth [32], Parti [39], and DALL E2 [13] demonstrating excellent effects in text-to-image generation.

However, there is still potential to improve the performance of these methods in professional domains such as interior design, especially in generating such designs for spaces of a specified decor and function.

This study proposes a new interior design diffusion model to generate better interior designs. The method includes a new decoration style dataset (i.e., IDSSA-64) and a new composite loss function (combining decoration style and spatial function as a loss). We retrain the diffusion model using this dataset and a new loss function. Thus, we obtain a model that can batch-generate interior designs of a specified decoration style and spatial function. This approach enhances the applicability of diffusion models in interior design, which enables the utilization of text-to-image generative design as a novel approach.

In regard to building the dataset, to solve the problem of the need for interior decoration style datasets, we collected more than 10,000 interior design images from several mainstream interior design websites associated with professional designers. We manually marked the decoration style and space function corresponding to each image. There are 8 categories of decoration styles and spatial functions, and 64 categories are formed by the pairwise combination of decoration style and spatial function. Thus, the "Interior Decoration Style and Spatial Attributes Dataset" (IDSSA-64) is established.

This study proposes a novel composite loss function for fine-tuning the model in the loss function construction stage. We add style prior preservation loss and style-aware reconstruct loss to the conventional loss function (Equation (1)) to form Equation (2). Style prior preservation loss retains the knowledge learned by the traditional model from big data, and style-aware reconstruct loss understands the decoration style and space function. These two losses constitute Equation (2) (i.e., the overall loss function). The new loss function enables the retrained diffusion model to retain the basic knowledge from big data and realize the decoration style and spatial function knowledge.

In the fine-tuning stage of the diffusion model, we utilized the decorative style and spatial function as the subject descriptors. We fine-tuned the diffusion model with tens of thousands of images in 64 categories in IDSSA-64. Specifically, these newly collected interior design data will be input into the diffusion model with a new compound loss function for training because the loss function includes the decoration style and spatial function loss items. The diffusion model will learn the decoration style during training, and it will learn the spatial function knowledge. Thus, it will generate an interior design with a specified decoration style and spatial function.

The basic diffusion model is given by using Equation (1):

$$\mathbb{E}_{\boldsymbol{Y},\boldsymbol{g},\boldsymbol{\epsilon},\boldsymbol{t}}[w_t||\hat{Y}_\theta(\alpha_t Y + \sigma_t \epsilon, \boldsymbol{g}) - Y||_2^2] \tag{1}$$

$\mathbb{E}$ in Equation (1) represents the expectation, meaning that all samples in the training set are optimized and averaged. $\hat{Y}_\theta$ is a diffusion model which reads a noisy image vector $\alpha_t Y + \sigma_t \epsilon$ and a text vector $\boldsymbol{g}$ and predicts a noise-free image. In simple terms, the conditional diffusion model is trained by reducing the loss of the loss function to denoise noisy pictures, while w_t is a weight parameter used to control different timestamps because the diffusion model will add noise to each training timestamp, and the w_t weights corresponding to pictures with different noise levels are also different.

The fine-tuning diffusion model is given by using Equation (2):

$$\mathbb{E}_{\boldsymbol{Y},\boldsymbol{g},\boldsymbol{\epsilon},\boldsymbol{\epsilon}',\boldsymbol{t}}[w_t||\hat{Y}_\theta(\alpha_t Y + \sigma_t \epsilon, \boldsymbol{g}) - Y||_2^2 + \lambda w_{t'}||\hat{Y}_\theta(\alpha_{t'} Y_{pr} + \sigma_{t'} \epsilon', \boldsymbol{g_{pr}}) - Y_{pr}||_2^2] \tag{2}$$

Equation (2) adds the style-aware reconstruct loss and style prior preservation loss to Equation (1) and uses interior design category information (i.e., decoration style and space function) as part of the loss function, where $\boldsymbol{g_{pr}}$ refers to the control vector with interior design category information and Y_{pr} relates to the data generated with the frozen pretrained diffusion model. The loss function encourages the model to retain the knowledge of the original model while learning the knowledge of the decoration style and space function so that the model can generate the specified decoration style and corresponding space.

At the stage where the user uses the fine-tuning model, the user only needs to input a text description of the decoration style and space function to generate the interior design. Figure 5 shows the difference in the design flow between the conventional design and the approach proposed in this study. Our method only needs to generate the design by inputting design descriptors into our fine-tuned diffusion model, thereby avoiding the traditional methods of collecting design reference diagrams, drawing 2D images, creating 3D models, and rendering images. The method proposed in this study optimizes the design process, improving efficiency and creativity.

In the design modification stage pertaining to the traditional design method, once the customer is unsatisfied with the design proposal, the entire design process (e.g., redrawing the 2D drawing, modifying the 3D model, and rerendering the rendering) must undergone again to complete the design modification. The whole modification cycle is prolonged, and real-time feedback cannot be obtained. In the proposed method, the designer should change only the generated prompt words according to the customer's needs and input the trained diffusion model. Thus, they can quickly obtain a new design scheme. Designers utilize the generated design plan to communicate with customers again, effectively reducing communication and enhancing decision-making efficiency.

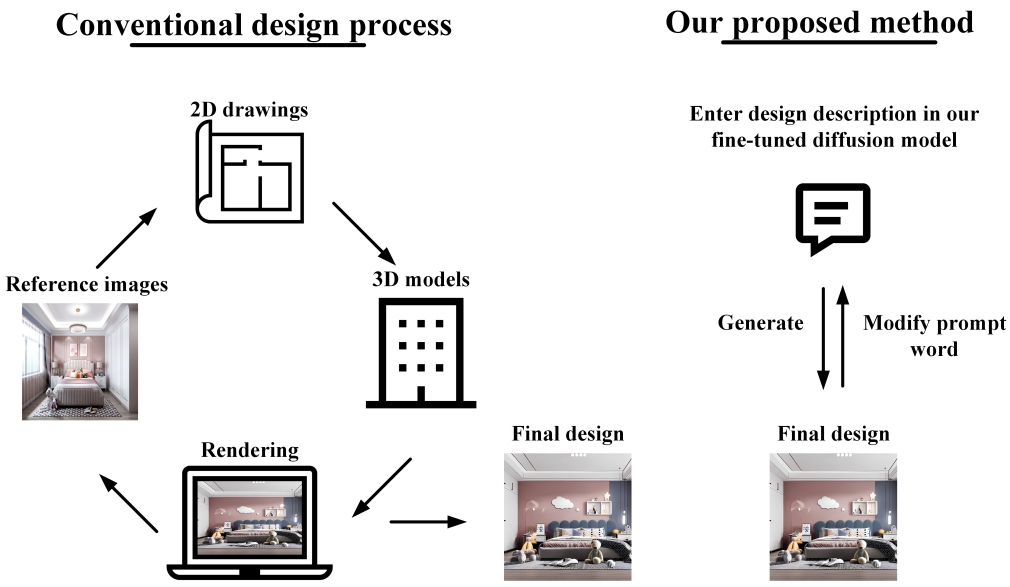


Figure 5. Comparison of different design methods. The conventional design process is complex and time-consuming. Our method generates an interior design through text descriptions of text decoration styles and space functions, thereby improving the efficiency of creative design.

4. Experiments and Results

4.1. Implementation Details

The fine-tuning diffusion model is based on PyTorch. The computer used in the experiment had a Windows 10 operating system with 32 GB of memory and a 16 GB graphics card. We set one million steps to the number of iterations and 72 h for each training. The preprocessing method automatically cuts the input image to a 512 × 512 resolution, sets the learning rate to 0.000002, sets the batch size to 8, and uses XFormers and FP16 to accelerate calculations.

4.2. Decorative Style Dataset

This study aims to mass-generate interior designs with specified decoration styles and space functions through text. To this end, this study first constructed a dataset, IDSSF-64, with descriptions of decoration styles and space functions, which were collected by professional designers from multiple interior design websites, and classified the collected

images. The designer marked two pieces of information on the image: the categories of decoration style and space function. Among them, there were eight categories for the decoration style and space function. Combining the two created 64 categories and finally formed the interior decoration style dataset IDSSF-64, with over 10,000 images. The decoration style categories included "Chinese style", "Nordic style", "Japanese style", "Modern style", "American style", "European style", "Mix style", and "Industrial style". The space function categories included "Living room", "Dining room", "Kitchen", "Study room", "Bedroom", "Kids room", "Bathroom", and "Entrance". The distribution of the number of pictures is shown in Table 1.

Table 1. Image distribution of each decorative style and space function corresponding to the IDSSF-64 dataset.

	Chinese Style	Nordic Style	Japanese Style	Modern Style	American Style	European Style	Mix Style	Industrial Style
Living room	309	730	54	774	67	525	232	231
Dining room	154	513	22	1149	24	294	130	74
Kitchen	28	28	12	99	20	18	56	17
Study room	90	37	21	186	19	62	42	25
Bedroom	327	607	20	668	47	357	162	115
Kids room	18	28	10	231	12	14	21	8
Bathroom	61	518	8	488	43	163	103	85
Entrance	37	67	14	329	16	59	32	22
Grand total	1024	2528	161	3924	248	1492	778	577

The IDSSF-64 dataset has the most modern-style images (3924 images) and the fewest Japanese-style images (161 images) under the decorative style category. However, within the scope of space functions, the living room has the most images (2922 images), and the children's room has the least (342 images). Figure 6 shows some high- and low-quality images in the training dataset. We used these annotated data to fine-tune the diffusion model and generate high-quality interior designs. In addition, the images and design quality generated by the fine-tuning of the diffusion model surpassed the training data, demonstrating the significant advantages of combining the fine-tuning method based on the basic diffusion model.



Figure 6. Partial image display of the IDSSF-64 dataset.

4.3. Visual Evaluation Metrics

Interior design evaluation is complex, and conventional automated image evaluation metrics are primarily utilized to measure image clarity, which is unsuitable for the proposed study [40–42]. Interior design should not only evaluate whether the image is clear; it should also consider whether the interior decoration style and space function of the generated image is rational. Because evaluating a space function and decoration style is subjective, it is necessary to assess the design manually. Therefore, this study invited several industry-based senior designers to discuss and create a series of evaluation indicators suitable for professional interior design. Thus, it comprehensively evaluates the quality of the generated interior design images. There were seven indicators, namely "Decoration Style", "Space Function", "Furniture Position", "Object Integrity", "Design Details", "Reality", and "Usability".

"Decoration Style" means that the generated interior design decoration style is consistent with the prompt word. "Space Function" means that the size of the generated space is appropriate, and the space function is compatible with the prompt word. "Furniture Position" means that the position of the generated furniture is reasonable. "Object Integrity" means that the generated object has no defects."Design Details" means that the generated image has rich details, such as realistic lighting and regular shapes for the furniture and walls. "Reality" means that the resulting image looks like a photo taken by a camera. "Usability" refers to the feasibility of directly using the generated design image, which includes aesthetic evaluation and functional evaluation, and it is a comprehensive evaluation index.

4.4. Visual Assessment

This study visually compares our diffusion model with other popular diffusion models (i.e., Disco Diffusion [43], Midjourney [21], DALL E2 [13], and Stable Diffusion [22]). Herein, we discuss the most mainstream and influential models. Midjourney has 14.5 million registered members, of which 1.1 million are active users [44]. DALL E2 has nearly 1 million active users [45], whereas Stable Diffusion has 10 million active users [46]. Disco Diffusion is an earlier model for generating artistic images, and it reflects the earlier diffusion models' ability to generate images [47].

We uniformly used "Chinese style study room with bookcases, tables and large windows, large angle, realistic, photo, high-definition" as the guide words to generate images, which are shown in Figure 7.

Figure 7 shows that Disco Diffusion [43] could not generate the interior design. The model needed help understanding the spatial-positional relationship of each component element in interior design, and the combined effect of each component element was inferior. Compared with Disco Diffusion [43], Midjourney [21] better understood the size of the space that the interior design should have had. The furniture was all in the right place, and the generated objects also had design details and a specific design aesthetic. But the designs generated by Midjourney [21] were painterly rather than photorealistic. The advantage of DALL E2 [13] is that the generated image style is highly realistic. The disadvantage is that most generated furniture needs to be completed, and the space size needs to be more reasonable. The advantage of Stable Diffusion [22] is that the generated image has details, and the decoration style is correct. The position of the generated object is also correct, but the disadvantage is that the empty size is unreasonable. For example, we needed to create a study room, and the generated space scale was closer to a library.

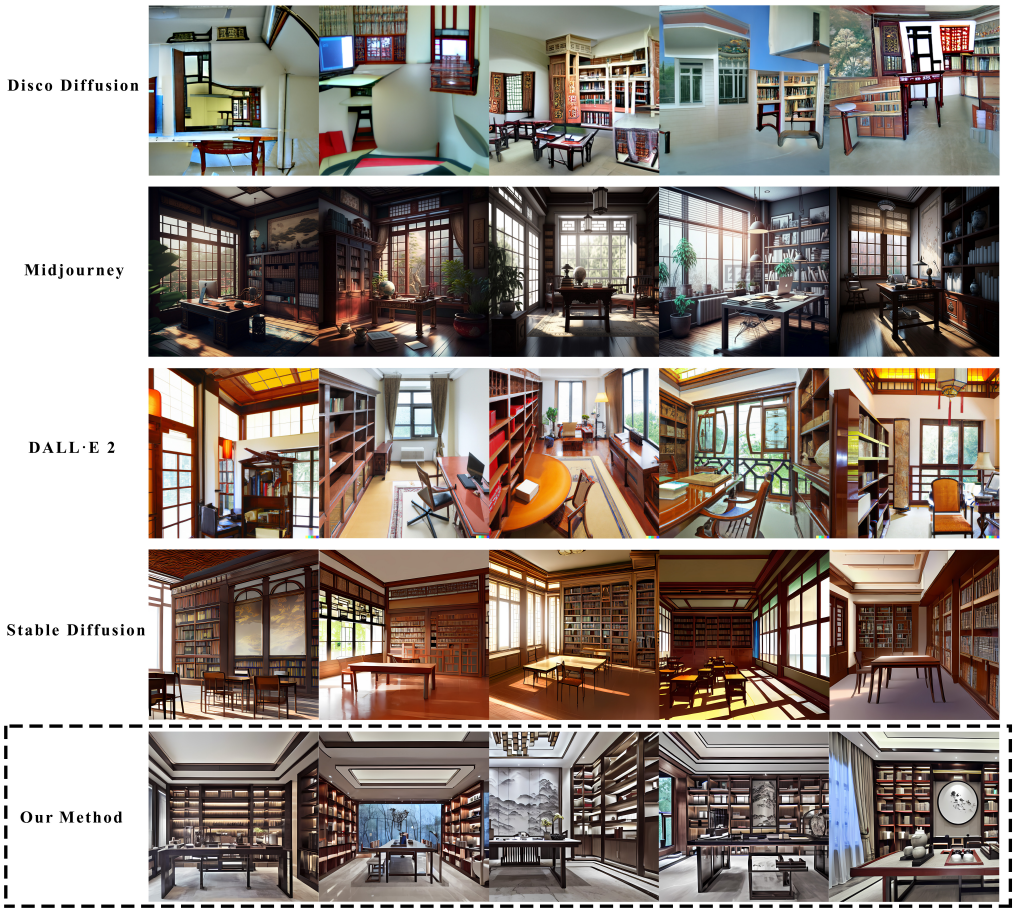


Figure 7. Comparison of other popular diffusion methods with our method for generative interior design images.

Compared with other methods, our method can generate an interior design with a specified decoration style and space function, and the design details are rich and realistic. The pros and cons of the above methods are shown in Table 2, which shows that the method proposed in this study was the best among all the tested methods, followed by Midjourney [21] and Stable Diffusion [22]. Designers cannot use DALL E2 [13] and Disco Diffusion [43] to generate interior designs.

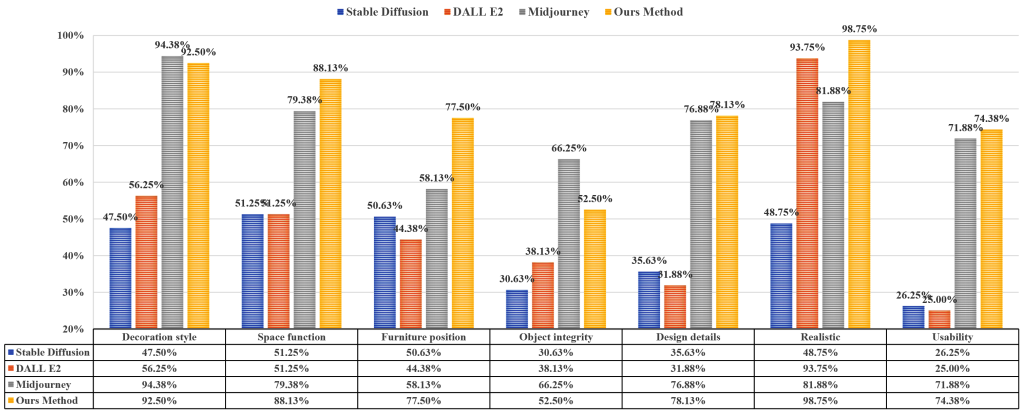
Table 2. Comparison of image effects generated by different diffusion models.

Model	Decoration Style	Space Function	Furniture Position	Object Integrity	Design Details	Realistic	Usability
Disco Diffusion						✓	
Midjourney	✓	✓	✓	✓	✓		✓
DALL E2	✓				✓	✓	
Stable Diffusion	✓		✓	✓	✓		✓
Our Method	✓	✓	✓	✓	✓	✓	✓

4.5. Quantitative Evaluation

Since Disco Diffusion [43] could not generate usable interior designs, we quantitatively evaluated the designs generated by the remaining four diffusion models (i.e.,

Midjourney [21], DALL E2 [13], Stable Diffusion [22], and our method) to create interior design images of 64 categories (eight decoration styles × eight space functions). Ten images were generated for each category for a total of 2560 images. We asked 10 professional designers to evaluate the above images. There were seven evaluation indicators, namely "Decorative Style", "Space Function", "Furniture Position", "Object Integrity", "Design Details", "Reality", and "Usability". The evaluation method consisted of evaluating whether the image met the evaluation index and adding one point if it did. Finally, the total score of each index was divided by the total number of images to obtain the average score, which was converted into a percentage to obtain the quantitative score of each model. The scores of the different diffusion models are shown in Figure 8:



	Decoration style	Space function	Furniture position	Object integrity	Design details	Realistic	Usability
Stable Diffusion	47.50%	51.25%	50.63%	30.63%	35.63%	48.75%	26.25%
DALL E2	56.25%	51.25%	44.38%	38.13%	31.88%	93.75%	25.00%
Midjourney	94.38%	79.38%	58.13%	66.25%	76.88%	81.88%	71.88%
Ours Method	92.50%	88.13%	77.50%	52.50%	78.13%	98.75%	74.38%

Figure 8. Different diffusion models generated quantitative assessments of interior design.

From Figure 8, it can be seen that there were considerable differences in the performance of these four models in the field of interior design. Our method and Midjourney [21] surpassed DALL E2 [13] and Stable Diffusion [22] in all indicators (except that Midjourney's performance in the realism indicators was not as good as that for DALL E2 and our method). Compared with other models, ours achieved the best results in the five indicators of space function, furniture location, design details, realism, and usability. Our model ranked second in both the decorative style and object integrity assessments, slightly behind Midjourney [21] but ahead of Stable Diffusion [22] and DALL E2 [13] by a wide margin. We believe that the reason why our model did not achieve the best results in decoration style and object integrity is that some decoration style data were challenging to collect, resulting in insufficient training of some categories and the generation of unsatisfactory images, (For example, the industrial-style children's room only had eight training images.) which can be solved by recollecting rare data and retraining the model.

To sum up, our method and Midjourney [21] are available for interior design generation (the usability exceeds 70%), and DALL E2 [13] and Stable Diffusion [22] are not available (that is, the usability is only about 25%). Compared with the interior design space generated by Midjourney [21], our method had a higher accuracy rate for interior design space functions, fewer examples of wrong furniture, and more design details and was more realistic. In addition, our method outperformed Midjourney by 2.5% on the usability index.

4.6. Different Decoration Styles' Design Generation

This section demonstrates our proposed method for rapidly generating interior designs with different decoration styles. Generating designs with varying types of decoration only requires a change in the subject words. Taking the generated study room as an example, we used words to describe a "study room with desks and bookcases" and added keywords trained by us, such as "Chinese style", "Nordic style", "Japanese style", "Modern style", and "American style". The resulting interior design is shown in Figure 9. The abscissa is the generated study room with different decoration styles, and the ordinate shows the effect of the room generated by three commonly used samplers. The fine-tuned model learned

the decoration style from the abscissa. Each decoration style had obvious differences. The positions and proportions of the generated objects were appropriate, and there were no obvious errors. We also found that the model generated small decorative objects such as books and plants, showing that our method can create designs with a large amount of detail, especially the generated design images with realistic reflection and refraction effects, which reached the level of ordinary renderings. From the ordinate point of view, even if the image decoration style generated by different samplers were changed, the image decoration style was very stable. We found that the image generated by the DDIM sampler was better than those from Euler and Euler a, the image generated by DDIM was more detailed, and the reflection and refraction effects were more realistic. While DDIM and Euler had incredibly similar generated images, there were subtle differences.



Figure 9. Comparison of interior design effects generated by different samplers.

4.7. Quick Modification of a Design

This section shows how to quickly modify a design by changing guide words and using the partial redraw feature. Figure 10 shows the effect of changing the guide word to produce a new design, and panel (a) shows the original image. Panel (b) transforms the cabinet on the right into a Chinese-style cabinet by replacing the guide word with "Chinese cabinet". Panel (c) adds the guide word "curtain" to add curtains to the design. Panel (d) replaces the guide word with "bathroom" to turn the original desk into a bathtub. Panel (e) changes the guide word to "Chinese style" to make the whole design a Chinese study.



Figure 10. Quickly modifying the design by changing the prompt word.

The above operations demonstrate the feasibility of rapidly modifying designs using text-guided fine-tuning of diffusion models. This approach allows for timely changes to local and global designs and modifies the image directly. After confirming that the design meets the requirements, the designer can directly produce the construction drawing without repeating the design process due to repeated revisions.

4.8. Generative Design Diversity

This section demonstrates the diversity of interior designs generated by our diffusion model. The prompt words were "Chinese-style study room with desk and bookcase". We set five random seeds to create different design images. The resulting image is shown in Figure 11. Figure 11 shows that the images maintained the same decoration style and space function but differed significantly in detail, indicating that the diffusion model trained in this study can generate diverse designs with specific decoration types and spatial functions. In particular, we note that the generated designs are creative. For example, the chest generated by Seed 1 is different from the standard bookcase. The bookcases were designed to be narrow, high, deep, and separated by thick wooden lines, which is rare in conventional designs. This method of generating diverse designs in batches can help designers select appropriate designs from generated works, find design inspiration, and accelerate the thinking process, thereby accelerating the generation of creative designs.



Figure 11. Different random seeds for the same prompt word generate diverse designs.

4.9. Generated Design Details Showcase

Figure 12 shows the Chinese-style study room generated by our diffusion model. This figure shows that our diffusion model can generate a structurally sound design with furniture in the right position and at the right size and no inappropriate objects generated. The picture creates many decorations, such as books, plants, and hanging paintings, which reflects the ability of the model to create design details. The resulting image was realistic, and the relationship between light and shadow was handled well. However, there is still room for improvement; that is, the details of the generated objects were not entirely correct. For example, the generated shape of the book was not a regular rectangle, and the thickness of the corners of the table needed to be more consistent. In general, the interior design generated using our diffusion model is already available, which can accelerate the generation of design proposals and the efficiency of design decision making.



Figure 12. An example of a Chinese-style study room generated using our trained diffusion model. The figure shows that the generated design has design details, and the generated decoration style and spatial function are consistent with the prompt words used (prompt word: Chinese-style study room, with desks and bookcases).

5. Discussion

This study demonstrates the effectiveness of the proposed approach through qualitative and quantitative research. For example, in the qualitative research section, through a visual comparison of the proposed fine-tuned diffusion method and other methods, it was proven that unlike other mainstream diffusion models, the fine-tuned diffusion model can quickly generate a high-precision interior design with diversity and a specified decoration style from end to end. In the quantitative research section, the quantified data prove that regarding many indicators (especially in designing the specified decoration style, space function, and usability), the proposed method is significantly superior to the other ones. For example, in comparing the spatial function indicators, the proposed method surpassed Midjourney's method by 8.75%, DALL E2, and the stable diffusion one by 36.88%, indicating the proposed method's effectiveness.

To change the traditional interior design linear workflow, this study utilized a text-guided diffusion model for rapid design acquisition, avoiding traditional manual modeling and rendering work. The diffusion models exhibited an absolute advantage in generating and modifying design efficiency. For example, it takes about a week to traditionally complete an interior design and a week to alter the design. Using the proposed method, running the diffusion model on a computer with 12 GB of video memory generates a design in about six seconds (10 designs per minute). If the design requires modification, then the designer must only change the prompt word to regenerate the design. The designers can choose from many different decoration styles generated by computers, which increases the possibility of designers obtaining creative designs. In summary, the proposed method can accelerate the design proposal generation and modification process, increase the likelihood of obtaining creative designs, and expedite the design decision-making process.

But the word-guided diffusion model has some drawbacks. Although the text-guided diffusion model can generate decorative styles and space functions, and the generated images contain rich design details, the diffusion model cannot control where the generated objects appear. At the same time, the generated design image cannot completely match the actual space, which leads to the need for the designer to modify the design according to the actual site size after using this method, and the customer quickly determines the scheme.

The application of diffusion models in interior design may exert profound implications. For designers, their role is redefined: They are no longer traditional design creators but instead transformed into design facilitators collaborating with AI. Designers should leverage their expertise and extensive experience to optimize and fine-tune the AI models' outputs.

Thus, they can meet clients' needs and expectations more effectively. This transformation also requires designers to continuously acquire and familiarize themselves with AI-related knowledge. For clients, the introduction of diffusion models enables them to be more profoundly and directly involved in the design process. They can observe the designs that are rapidly generated and adjusted by designers using AI in real time, which effectively enhances the communication efficiency and design quality. For the entire interior design industry, it is apparent that the emergence of diffusion models immensely facilitates design efficiency and innovation, and thus bulk design generation becomes more convenient.

We propose that inherent ethical risks are associated with the utilization of AI to generate designs. AI training is highly dependent on a substantial amount of internet-obtained data. Although we affirm that publicly available data collection has crucially facilitated the development of the computer field and reiterate the fundamental rationale of this practice, ensuring personal privacy during data collection is imperative. To address this issue, we propose the creation of appropriate legislation or the implementation of data usage agreements by data owners upon data upload, which are potentially viable approaches. In regard to which data can be freely utilized and which data require the prior application for usage permissions, the aforementioned measures would enable more apparent differentiation for data users. Therefore, we can optimally safeguard the security of personal privacy.

6. Conclusions

Due to traditional interior design's inefficiency and lack of creativity, this study proposes improving the above problems by fine-tuning the diffusion model. This study built a unique dataset of interior decoration styles to support the development of interior design diffusion models. This study used this dataset to fine-tune existing diffusion models. The fine-tuned model can generate batches of interior design images that conform to a decoration style and space function. Compared with the traditional design method, our method automatically turns the tedious design process of conventional design into computer-generated design. It can modify the design image nearly in real time when guided by the text, thus improving efficiency. At the same time, the designs generated in batches provide designers with rich creative inspiration and solves the problem of designers needing more creativity. Experiments showed that the method proposed in this study is a new creative design method that can efficiently generate diverse interior designs, significantly helping to improve efficiency and creativity.

Furthermore, we note that the diffusion model-based generative design exhibits certain limitations and poses some social issues:

1. It impacts the career prospects of designers. Owing to increased productivity, those who master these techniques first will exert an elimination effect on other designers who do not.
2. The generative interior design technology may solidify a design style and affect people's aesthetic tendencies. Diffusion models generate designs by learning from datasets, which are the source of innovation. Consider a scenario where many designers utilize the same model on a large scale without adding new training data. In the aforementioned scenario, the designs generated by the model will flood the network, and people will easily access these designs. Thus, their aesthetics will be affected. Therefore, someone must constantly create new images to supplement the model's training.
3. Because the development of generative image technology will weaken people's ability to judge the authenticity of images, people will become more vulnerable to deception.
4. Maintaining cultural diversity is a challenge. Although diffusion models usually require massive datasets to complete training, these training datasets may need more cultural diversity. Activity with these data could lead to computer-generated imagery lacking diversity and inclusivity, affecting wider audiences.

5. Quantitative evaluation indicators are difficult to establish. Although experts have developed the current quantitative evaluation index to measure the design quality comprehensively, further enhancements are imperative. However, it is challenging to formulate more comprehensive quantitative indicators that can reflect subjective feelings.

We note the following for future research:

1. This study focused only on the more mainstream interior design styles and did not collect and train data on niche decoration styles. Collecting more types of indoor datasets can enhance the generation effect of interior design and serve a wider group of people.
2. Relying on only text-guided image generation cannot generate precisely controlled designs. Adding more design constraints to the generation process, such as using multimodal control image generation, is a future research direction.
3. Diffusion models are not yet capable of generating high-quality 3D designs. In the future, it will be a better way to directly generate three-mode interior design models or videos, allowing users to feel the design more intuitively.
4. Since interior design evaluation includes subjective evaluation, such as decoration style and space function evaluation, it is highly dependent on manual scoring, and there is an urgent need for automatic quantitative interior method assessment.
5. The rendering-based interaction between designers and users is limited, and establishing a synchronous design platform may enable users to offer real-time feedback on designers' designs.

Author Contributions: Conceptualization, J.C. and Z.S.; data curation, J.C. and Z.S.; formal analysis, J.C.; funding acquisition, B.H.; investigation, J.C.; methodology, J.C.; project administration, B.H.; resources, B.H.; software, J.C.; supervision, B.H.; validation, J.C. and B.H.; visualization, J.C.; writing—original draft, J.C. and Z.S.; writing—review and editing, J.C. and Z.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Integrated Design of Smart Park Based on Digital City, Soochow Social Science (No: Y2023LX013).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The labeled dataset used to support the findings of this study is available from the authors upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, Y.; Liang, C.; Huai, N.; Chen, J.; Zhang, C. A Survey of Personalized Interior Design. In *Computer Graphics Forum*; Wiley Online Library: Hoboken, NJ, USA, 2023. [CrossRef]
2. Ashour, M.; Mahdiyar, A.; Haron, S.H. A Comprehensive Review of Deterrents to the Practice of Sustainable Interior Architecture and Design. *Sustainability* **2021**, *13*, 10403. [CrossRef]
3. Bao, Z.; Laovisutthichai, V.; Tan, T.; Wang, Q.; Lu, W. Design for manufacture and assembly (DfMA) enablers for offsite interior design and construction. *Build. Res. Inf.* **2022**, *50*, 325–338. [CrossRef]
4. Karan, E.; Asgari, S.; Rashidi, A. A markov decision process workflow for automating interior design. *KSCE J. Civ. Eng.* **2021**, *25*, 3199–3212. [CrossRef]
5. Park, B.H.; Hyun, K.H. Analysis of pairings of colors and materials of furnishings in interior design with a data-driven framework. *J. Comput. Des. Eng.* **2022**, *9*, 2419–2438. [CrossRef]
6. Sinha, M.; Fukey, L.N. Sustainable Interior Designing in the 21st Century—A Review. *ECS Trans.* **2022**, *107*, 6801. [CrossRef]
7. Delgado, J.M.D.; Oyedele, L.; Ajayi, A.; Akanbi, L.; Akinade, O.; Bilal, M.; Owolabi, H. Robotics and automated systems in construction: Understanding industry-specific challenges for adoption. *J. Build. Eng.* **2019**, *26*, 100868. [CrossRef]
8. Sohl-Dickstein, J.; Weiss, E.; Maheswaranathan, N.; Ganguli, S. Deep unsupervised learning using nonequilibrium thermodynamics. In Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 6–11 July 2015; Bach, F., Blei, D., Eds.; PMLR: New York, NY, USA, 2015; Volume 37, pp. 2256–2265. [CrossRef]

9. Croitoru, F.A.; Hondru, V.; Ionescu, R.T.; Shah, M. Diffusion models in vision: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, 1–20. [CrossRef] [PubMed]
10. Liu, X.; Park, D.H.; Azadi, S.; Zhang, G.; Chopikyan, A.; Hu, Y.; Shi, H.; Rohrbach, A.; Darrell, T. More control for free! image synthesis with semantic diffusion guidance. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 2–7 January 2023; pp. 289–299. [CrossRef]
11. Ho, J.; Salimans, T. Classifier-Free Diffusion Guidance. *arXiv* **2021**, arXiv:2207.12598.
12. Nichol, A.Q.; Dhariwal, P. Improved denoising diffusion probabilistic models. In Proceedings of the International Conference on Machine Learning, Virtual Event, 18–24 July 2021; Meila, M., Zhang, T., Eds.; PMLR: New York, NY, USA, 2021; Volume 139, pp. 8162–8171. [CrossRef]
13. Ramesh, A.; Dhariwal, P.; Nichol, A.; Chu, C.; Chen, M. Hierarchical text-conditional image generation with clip latents. *arXiv* **2022**, arXiv:2204.06125. [CrossRef]
14. Saharia, C.; Chan, W.; Saxena, S.; Li, L.; Whang, J.; Denton, E.L.; Ghasemipour, K.; Gontijo Lopes, R.; Karagol Ayan, B.; Salimans, T.; et al. Photorealistic text-to-image diffusion models with deep language understanding. In *Advances in Neural Information Processing Systems*; Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., Oh, A., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2022, Volume 35, pp. 36479–36494. [CrossRef]
15. Song, Y.; Ermon, S. Improved techniques for training score-based generative models. In *Advances in Neural Information Processing Systems*; Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., Lin, H., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2020; Volume 33, pp. 12438–12448. [CrossRef]
16. Song, Y.; Ermon, S. Generative modeling by estimating gradients of the data distribution. In *Advances in Neural Information Processing Systems*; Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Garnett, R., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2019; Volume 32. [CrossRef]
17. Gu, S.; Chen, D.; Bao, J.; Wen, F.; Zhang, B.; Chen, D.; Yuan, L.; Guo, B. Vector quantized diffusion model for text-to-image synthesis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 10696–10706. [CrossRef]
18. Nichol, A.Q.; Dhariwal, P.; Ramesh, A.; Shyam, P.; Mishkin, P.; Mcgrew, B.; Sutskever, I.; Chen, M. GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models. In Proceedings of the International Conference on Machine Learning, PMLR, Baltimore, ML, USA, 17–23 July 2022; pp. 16784–16804. [CrossRef]
19. Kawar, B.; Zada, S.; Lang, O.; Tov, O.; Chang, H.; Dekel, T.; Mosseri, I.; Irani, M. Imagic: Text-based real image editing with diffusion models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 18–22 June 2023; pp. 6007–6017. [CrossRef]
20. Avrahami, O.; Lischinski, D.; Fried, O. Blended diffusion for text-driven editing of natural images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 18208–18218. [CrossRef]
21. Borji, A. Generated faces in the wild: Quantitative comparison of stable diffusion, midjourney and dall-e 2. *arXiv* **2022**, arXiv:2204.06125. [CrossRef]
22. Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; Ommer, B. High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 10684–10695. [CrossRef]
23. Song, J.; Meng, C.; Ermon, S. Denoising Diffusion Implicit Models. In Proceedings of the International Conference on Learning Representations, Addis Ababa, Ethiopia, 26–30 April 2020. [CrossRef]
24. Jolicoeur-Martineau, A.; Piché-Taillefer, R.; Mitliagkas, I.; des Combes, R.T. Adversarial score matching and improved sampling for image generation. In Proceedings of the International Conference on Learning Representations, Virtual Event, 3–7 May 2021. [CrossRef]
25. Dhariwal, P.; Nichol, A. Diffusion models beat gans on image synthesis. In *Advances in Neural Information Processing Systems*; Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., Vaughan, J.W., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2021; Volume 34, pp. 8780–8794. [CrossRef]
26. Ding, M.; Yang, Z.; Hong, W.; Zheng, W.; Zhou, C.; Yin, D.; Lin, J.; Zou, X.; Shao, Z.; Yang, H.; et al. Cogview: Mastering text-to-image generation via transformers. In *Advances in Neural Information Processing Systems*; Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., Vaughan, J.W., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2021; Volume 34, pp. 19822–19835. [CrossRef]
27. Gafni, O.; Polyak, A.; Ashual, O.; Sheynin, S.; Parikh, D.; Taigman, Y. Make-a-scene: Scene-based text-to-image generation with human priors. In Proceedings of the Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, 23–27 October 2022; Proceedings, Part XV; Springer: Cham, Switzerland, 2022; pp. 89–106. [CrossRef]
28. Gal, R.; Alaluf, Y.; Atzmon, Y.; Patashnik, O.; Bermano, A.H.; Chechik, G.; Cohen-Or, D. An image is worth one word: Personalizing text-to-image generation using textual inversion. *arXiv* **2022**, arXiv:2208.01618. [CrossRef]
29. Choi, J.; Kim, S.; Jeong, Y.; Gwon, Y.; Yoon, S. ILVR: Conditioning Method for Denoising Diffusion Probabilistic Models. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, BC, Canada, 11–17 October 2021; pp. 14347–14356. [CrossRef]

30. Von Oswald, J.; Henning, C.; Grewe, B.F.; Sacramento, J. Continual learning with hypernetworks. In Proceedings of the 8th International Conference on Learning Representations (ICLR 2020) (Virtual), Addis Ababa, Ethiopia, 26–30 April 2020. [CrossRef]
31. Hu, E.J.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; Chen, W. LoRA: Low-Rank Adaptation of Large Language Models. In Proceedings of the International Conference on Learning Representations, Virtual Event, 25–29 April 2022. [CrossRef]
32. Ruiz, N.; Li, Y.; Jampani, V.; Pritch, Y.; Rubinstein, M.; Aberman, K. Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 18–22 June 2023; pp. 22500–22510. [CrossRef]
33. Lee, J.; Cho, K.; Kiela, D. Countering Language Drift via Visual Grounding. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 3–7 November 2019; Association for Computational Linguistics: Hong Kong, China, 2019; pp. 4385–4395. . [CrossRef]
34. Lu, Y.; Singhal, S.; Strub, F.; Courville, A.; Pietquin, O. Countering language drift with seeded iterated learning. In Proceedings of the International Conference on Machine Learning, Virtual Event, 13–18 July 2020; Daumé, H., III, Singh, A., Eds.; PMLR: New York, NY, USA, 2020; Volume 119, pp. 6437–6447. [CrossRef]
35. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Li, F. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255. [CrossRef]
36. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft coco: Common objects in context. In Proceedings of the Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, 6–12 September 2014; Proceedings, Part V 13; Springer: Cham, Switzerland, 2014, pp. 740–755. [CrossRef]
37. Jiang, Y.; Huang, Z.; Pan, X.; Loy, C.C.; Liu, Z. Talk-to-edit: Fine-grained facial editing via dialog. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 13799–13808. [CrossRef]
38. Ding, M.; Zheng, W.; Hong, W.; Tang, J. CogView2: Faster and Better Text-to-Image Generation via Hierarchical Transformers. In Proceedings of the Advances in Neural Information Processing Systems, New Orleans, LA, USA, 28 Novermber–3 December 2022. [CrossRef]
39. Yu, J.; Xu, Y.; Koh, J.Y.; Luong, T.; Baid, G.; Wang, Z.; Vasudevan, V.; Ku, A.; Yang, Y.; Ayan, B.K.; et al. Scaling Autoregressive Models for Content-Rich Text-to-Image Generation. *Trans. Mach. Learn. Res.* **2022**, 1–53. [CrossRef]
40. Sara, U.; Akter, M.; Uddin, M.S. Image quality assessment through FSIM, SSIM, MSE and PSNR—A comparative study. *J. Comput. Commun.* **2019**, *7*, 8–18. [CrossRef]
41. Yang, H.H.; Yang, C.H.H.; Tsai, Y.C.J. Y-net: Multi-scale feature aggregation network with wavelet structure similarity loss function for single image dehazing. In Proceedings of the ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 2628–2632. [CrossRef]
42. Ding, K.; Ma, K.; Wang, S.; Simoncelli, E.P. Image quality assessment: Unifying structure and texture similarity. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *44*, 2567–2581. [CrossRef]
43. Lyu, Y.; Wang, X.; Lin, R.; Wu, J. Communication in Human–AI Co-Creation: Perceptual Analysis of Paintings Generated by Text-to-Image System. *Appl. Sci.* **2022**, *12*, 11312. [CrossRef]
44. Wilson, A. Midjourney Statistics: Users, Polls, & Growth [June 2023]. Available online: https://approachableai.com/midjourney-statistics/ (accessed on 31 May 2023).
45. Bastian, M. DALL-E 2 Has More Than One Million Users, New Feature Released. Available online: https://the-decoder.com/dall-e-2-has-one-million-users-new-feature-rolls-out/ (accessed on 1 September 2022).
46. Kenrick Cai, A.K. Six Things You Didn't Know about ChatGPT, Stable Diffusion and the Future of Generative AI. 2023. Available online: https://www.forbes.com/sites/kenrickcai/2023/02/02/things-you-didnt-know-chatgpt-stable-diffusion-generative-ai/?sh=54fa6997b5e3 (accessed on 2 February 2023).
47. He, C. Stable Diffusion vs Disco Diffusion. 2022. Available online: https://chengh.medium.com/stable-diffusion-vs-disco-diffusion-99e3e8957c0d (accessed on 25 August 2022).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.