International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Room Interior Generation Using Stable Diffusion Models

Shubh Sahu[1], Aryan Ghogare[2], Utkarsh Sharma[3], Shreya Verma[4], Om Patil[5]

[1,2,3,4,5]Dept. of Computer Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University Pune, India

Abstract
The present paper puts forward an innovative method for the fine tuning of Stable Diffusion with the use of Dreambooth, which is a very fast technique in the image generation task. In particular, our approach covers the topic of room interior generation. It offers a quick option to instruct the algorithm with the most recent concepts, without the need to retrain it. Using a shortened list of instance and class prompts as examples we introduce a revolutionary training pipeline where class and instance information interacts to order a model's learning process. By conducting elaborate trials where we demonstrate that our methodology out-perform the competitors in the realistic visualization, which is based on the given theme. We use an extensive evaluation process to prove the effectiveness of the method on many datasets, which can ensure its generalization capacity to unseen various layouts of the rooms and their interior design. On one hand, precise ablation studies are carried out in order to evaluate the influence of the distinguished components in the given model. This paper shows the encompassing work of Dreambooth as a tool of choice for the personalized room interior synthesis. The new possibilities to fine-tune the generative models are also becoming a subject for futurological research in the interior design area.

Keywords: Interior Modification, Stable Diffusion, DreamBooth

I. INTRODUCTION

The recently emergent phenomena of personalization in the visual aspect of room space across the disciplines of interior design, architecture, VR and so on is not only in demand but adored by today's people due to the rapid evolution of the modern world. The ability of creating charming room surroundings, extraordinary work spots, innovative virtual reality environments and quickly generating the room interiors that satisfy the users preferences without wasting their time is crucial.

Conventionally, room space development has been a tedious process that has been dominated by manual labor and thus was human resource intensive and time consuming. Thus, it not only limits creativity but may also be a cause of challenges when one is creating user-friendly systems that anticipate and meet the changing needs of customers while they scale. Beyond that, the vast amount of detail possibilities to explore and as a result impact the speed to evaluate and create different designs what is difficult to do in an effective manner.

Developing in parallel with the technology boom are tools like Stable Diffusion and Dreambooth which have brought about a much needed shift in how room interiors are artistically rendered to incorporate the occupants' personality and comfort. These new-generation solutions based on the deep learning and generative models make the formerly lengthy processes done by humans more efficient, doable by different specialists and offer more creative choices.

Stable Diffusion: Detects of the Stable Diffusion are the usage of the diffusion models to produce versatile and realistic images for giving conditioning data. Different from the other generative methods, diffusion models don't deal with mode collapse and imprecision, therefore, achieve an impressive level of visual output quality. Skillful, in that it's based on massive training data of numerous rooms, the Stable Diffusion program learns to put into practice the intricate patterns and the finer details, bringing about pictures that are novel and appealing aesthetically speaking with unbelievable accuracy.

Dreambooth: Not stopping at where Stable Diffusion does, Dreambooth makes it possible to bring the idea of the perfect image in stability by tweaking its strategy of learning and personalization. A hands-on approach reduces the need for extensive training or complex algorithms. By giving just a few pictures as an input, Dreambooth empowers users to develop new ideas and styles to the Stable Diffusion model, making it bit more customizable and at the same time adjustable to their own specific preferences and demands. Whether through the implementation of specific architectural features, the advent of various color combinations, or the exploration of avant-garde patterns, Dreambooth provides immaculate levels of personalization and hands-on control over the creation or reconstruction process of a room.

Addressing the Need: We combined forces of both Stable Diffusion and Dreambooth which provided us with a suitable and customized room interior solution to one of the burning issues – the efficient and customized room interior generation. Through fusion of these technologies with AI and machine learning, artists and creators can better automate their design workflows, ameliorate nature of outputs and also make it more user friendly and client specific. The scale of these real world tools for the interior design and the architecture along with the immersive virtual worlds of the gaming and the entertainment can be seen as nothing than just transformative.

As these technologies mature and show their limitlessness in terms of their abilities and specific applications, it is evident that the possibilities of generating rooms in the future are limitless, limited only by our imagination. In the future paragraphs we will focus on the methodology, the results as well as consequences of stable-diffusion and dreambooth as a versatile tool to finally change the way we perceive and create room interiors in the digital era.

II. LITERTURE SURVEY

2.1 Related Publications

[1] ***Adding Conditional Control to Text-to-Image Diffusion Models by Lvmin Zhang, Anyi Rao, and Maneesh Agrawala*** - The publication presents Stable diffusion and ControlNet, a novel neural network structure enhancing spatial composition control in image generation beyond text prompts. Extensive experiments demonstrate ControlNet's ability to interpret diverse input conditioning images and produce high-quality results, showcasing robustness with limited datasets and scalability. Visual comparisons with previous methods illustrate ControlNet's superior performance. Its transferability to community models without retraining highlights its broad applicability, making it a promising advancement in text-to-image generation.

[2] TEXT TO IMAGE GENERATION USING STABLE DIFFUSION by Divyanshu Mataghare, Shailendra S. Aote, Ramchand Hablani - The proposed Latent Diffusion Models in this (LDMs) can generate high-quality images from text descriptions while requiring significantly less computational resources compared to pixel-based diffusion models. This is achieved by leveraging a perceptually-compressed latent space that retains the essential semantic and conceptual information while discarding high-frequency details. The LDMs inherit the image-specific inductive biases of the UNet architecture, making them well-suited for spatially structured data without the need for high compression levels that compromise quality.

[3] Editing in Style: Uncovering the Local Semantics of GANs by Edo Collins, Raja Bala, Bob Price, Sabine Ssstrunk - Firstly, it provides crucial insights into the structure of hidden activations in StyleGAN, revealing disentangled representations essential for effective editing techniques. Secondly, the paper introduces a novel editing method leveraging these representations for localized changes, achieving high photorealism without complex spatial processing. Lastly, through experimental validation and discussion of potential applications, including extensions to real image editing and incorporation into adversarial training, it demonstrates the method's effectiveness and sets the stage for future advancements in the field.

[4] Image Style Transfer Using Convolutional Neural Networks by Leon A. Gatys, Alexander S. Ecker, Alexander S. Ecker - The paper "Image Style Transfer Using Convolutional Neural Networks" presents a novel approach to image style transfer by utilizing CNNs to separate and recombine image content and style, offering insights into deep image representations and showcasing the Neural Algorithm of Artistic Style for creating high-quality images. It contributes to understanding advanced image processing techniques by demonstrating the potential of CNNs for transferring image style and highlighting technical advancements in the field, making it a significant contribution to image processing with valuable implications for future applications.

[5] Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization by Xun Huang, Serge Belongie - The paper introduces Adaptive Instance Normalization (AdaIN), enhancing real-time arbitrary style transfer by aligning content and style statistics. This method offers abundant user controls, including content-style trade-off and style interpolation, while demonstrating strong generalization ability. AdaIN enables flexible controls at runtime and outperforms other methods in speed, making it a novel and effective approach to style transfer.

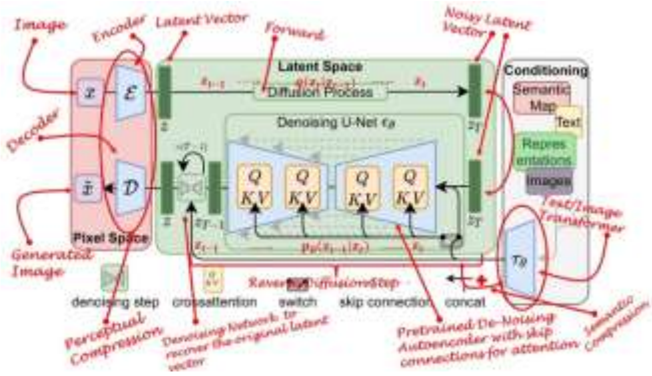
[6] Decorating Your Own Bedroom: Locally Controlling Image Generation with Generative Adversarial Networks by Zhang, C., Xu, Y., & Shen, Y. (2021). - The paper introduces LoGAN, a method for locally editing output images using GANs. It addresses the challenge of controlling image generation with pre-trained GANs by utilizing content modulation, style modulation, and a priority mask. This enables precise control of intermediate feature maps for local editing. LoGAN is effective in re-decorating bedrooms with customized furniture and styles, showcasing its versatility in image editing. The paper discusses advancements in GANs for image synthesis and presents experiments demonstrating LoGAN's application for local image editing and customization.

[7] mage2StyleGAN: How to Embed Images Into the StyleGAN Latent Space? Rameen Abdal,Yipeng Qin,Peter Wonka - The document introduces an efficient algorithm for embedding images into the latent space of StyleGAN, enabling semantic image editing operations such as morphing, style transfer, and expression transfer. It provides insights into the structure and capabilities of the StyleGAN latent space, including the diversity of embedded images, the meaningfulness of embedding, and the influence of the initial latent code. Through experiments, the algorithm's effectiveness, particularly on human faces, is demonstrated. Limitations regarding algorithm robustness, noise channels, and ImageNet-based perceptual loss are addressed, alongside potential applications in image editing operations.

III. METHODOLOGY

This research aimed to develop a system capable of modifying room images based on user-provided text descriptions. We leveraged the power of Stable Diffusion, specifically its DreamBooth implementation, to achieve this with surprising efficiency. Unlike traditional methods requiring vast image datasets, our approach thrives on a minimal training set – utilizing only 3-4 images of the target room.

Fig 1: Stable Diffusion Model



The core of our methodology lies in DreamBooth, a fine-tuning technique for Stable Diffusion. By feeding it a small collection of room images, we essentially create a customized diffusion model specifically attuned to that particular space. This allows the model to understand the room's unique characteristics, like its layout, furniture styles, and overall ambiance.

Following this training phase, textual prompts come into play. Users can provide descriptions of desired modifications – perhaps adding a new couch, changing the wall paint color, or incorporating specific decorative elements. The fine-tuned DreamBooth model then interprets these prompts and generates new images reflecting the user's vision, all while maintaining consistency with the original room's visual style captured in the training data.

In essence, our methodology combines the text-to-image generation capabilities of Stable Diffusion with the room-specific customization enabled by DreamBooth. This allows for efficient and user-friendly image editing with minimal training data requirements.

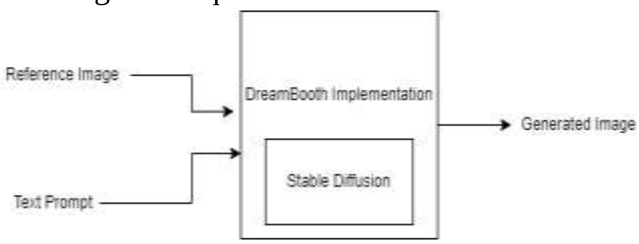


Fig 2: System Architecture of the Project

IV. APPLICATIONS

A. *Interior Design*: It is useful to understand the fundamental idea for the case of the design domain as well where the analytics of Stable Diffusion and Dreambooth offers the way the designers can change how the houses look like.Which this raises the designers can also achieve making room layouts which are both personalized and aesthetic, and do it fast due to the technological tools.Such plans may have the capability to materially alter the final compositions.

B. *Architecture*: The architects have great prospects in this technology as they can no longer survey the houses in 2 D and specify the screen configuration, room layout and the design while doing it.

C. *Virtual Reality (VR):* The goal of the course is to be able to vertically construct room interiors in the VR environment to have a wider range of purposes like gaming, simulations and virtual tour around the house property.
D. *Entertainment Industry*: Newer technologies in the entertainment field make the reality of the visuals truly-to-life and interactive for the audience, through movies, TV shows as well as other media formats.

V. RESULTS AND DISCUSSIONS

A. *Personalized Room Interiors*: The advantage of merging Stable Diffusion with Dreambooth is the opportunity for users to make highly customized digital-sided rooms that display the personal preferences of everybody's tastes. It helps customers to look for the information they need and be consistent and relevant results.
B. *Efficiency and Time Savings:* Automation of room interiors generation process makes it much faster and designers and architects may have more time in their work (not any more necessity to spend days/weeks for the same process).
C. *Enhanced Creativity:* This approach is partly due to the fact that a combination of technologies gives designers a chance to express themselves by creating different products and experimenting with colors and patterns in various ways, furthering design creativity and innovation in room interior design.
D. *Improved User Experience:* Finally, customers can see their designs quickly and very clearly, as well as get a feel of using them, which increases the chance of them selecting a choice and implementing it accordingly.

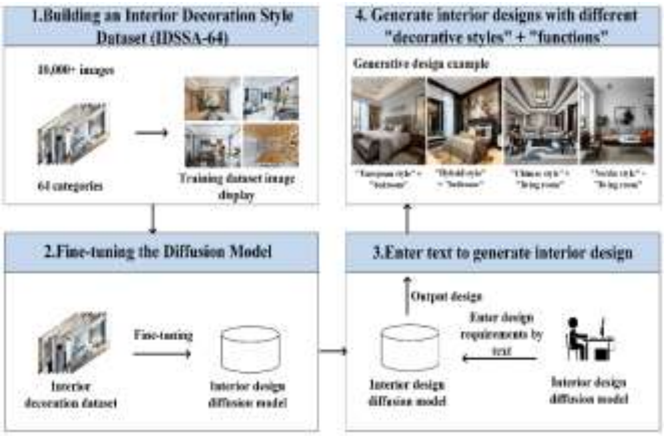


Fig 3: Text - Based Room Interior Generation

VI. POTENTIAL ISSUES

There has been significant focus on the use of computational strategies for the synthesis of room interior designs lately which provides with the room interior designers, architects and homeowners the new ways of doing so. Here we investigate potential issues associated with generating room interior designs using two prominent generative mode+6ls: the Stable Diffusion Model and Generative Adversarial Networks (GANs). Through the performance of comparative studies, the intention is to highlight the competencies and shortfalls of every method that is utilized in computational interior design. In doing so, future investigations in this area will be guided effectively.

A. Limited Diversity in Stable diffusion model

1. The Stable Diffusion Model has the potential problem of reproducing stable and consistent interior design of rooms. Based on the fact that it is built upon sequential sampling and diffusion processes, the model may be difficult to depict all the designs perfectly making the range of outputs to be limited.There is a possibility that diffusion models might be imbued with the nature that makes their generated designs to look alike, or uniquely similar, and may limit creativity and novelty in the synthetic decor of the room.

B. Lack of spatial coherence

1. The GAN model has proved itself an extraordinary tool that is capable of generating realistic images not only in different domains, but also faces the issue of inconsistency and incoherence while creating three-dimensional room interiors.
2. The adversarial training process in GANs can not account completely for the intricate spatial interactions as well as the contextual clues necessary to create cohesive room layouts hence can give rise to anomalies such as distorted or impossible placements of furniture, lighting and architectural elements.

C. Data quality and Bias

1. Generators that are based on Stable Diffusion model and GAN model heavily depend on the quality and diversity of training data. Problems of bias towards certain datasets, absence of many architectural styles, and not enough replacements of the room configurations can diminish the precision and originality of machine-generated interior designs.

VII. EVALUATING EFFECTIVENESS

Fig 4: Room Image : room123

Fig 5: Image generated after text prompt - Generate an image on room123 and add a shelf on wall with some items like plants on them

Fig 6: Image generated after text prompt - Generate an image of room 123 with the bed and a beach view painting hanging on the wall

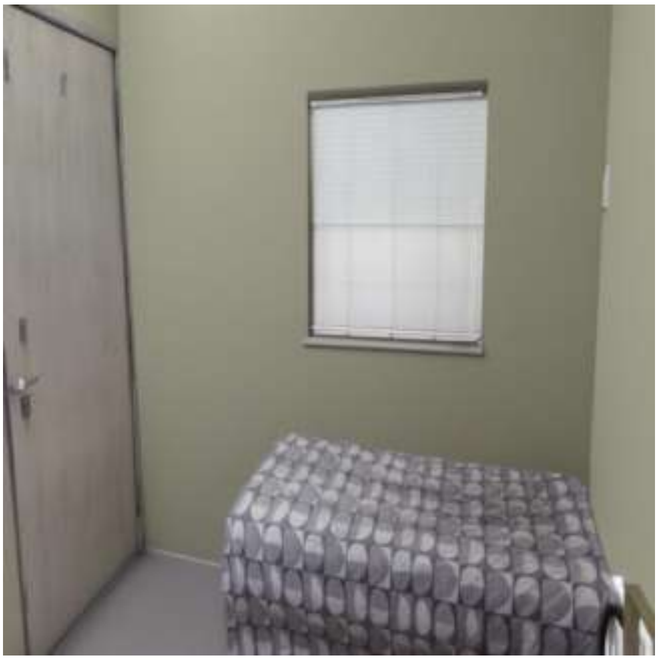


Fig 7: Image generated after text prompt – Generate an image of room 123 with a big window in the center along with the bed

Fig 8: Image generated after text prompt - Generate an image of room 123 with a bed and add an abstract painting hanging on the wall

An appropriate level of evaluation is very critical when it comes to the use of the pretrained set of stable diffusion tools in AI - generated art designs when it comes to the room's inserted designs. This research investigate the potential of using pretrained models to address the issue of realism and variety in VR interior designs by exploring their capability to show complicated space structure, light nuances and artistic sensitivities.

This study intends to evaluate systematically the machine produced outputs : an analysis that bypasses human concepts about their mistakes and regulations. Furthermore, in the study, the target is to assess the effectiveness and ease of the diffusion models used as automation processes in interior design practices. Through the process, the learning points which are extracted not only assists in the foundation of artificial intelligence-oriented design methodologies but also provides adequate advice for architects and designers who strive to explore new smart ways of designing spaces.

VIII. FUTURE PROSPECTS

1. Advanced Fine-Tuning Techniques:

The possibility of continuous investigation and subsequent development in fine-tuning of Stable Diffusion and Dream Booth techniques may also result in higher fidelity and accuracy for image generation exercises. Whether the focus is on the development of cutting-edge techniques for this purpose or in the exploration of different applications in which these models can be applied, the investigation will be carried out.

2. Personalized Interior Design Solutions:

The immediate programming of smart algorithms with newly introduced topics without intervention as a middleman roles the development of personalized interior design solutions. The next research aspect may involve the development of user-friendly and simple interfaces that will help users specify their preferences and plans receiving responsive suggestions on the room interior.

3. Integration with Augmented Reality (AR) and Virtual Reality (VR):

The tools and technology can be combined to showcase how concrete applications of generative models for room interior devices can help transform the interior design industry for the better. The users are able to see the exact layout of the virtual room and can also inform them in real time and make the informed decision before starting the physical changes.

4. Cross-Domain Applications:

The methods discussed in this work on the other hand could be also applied to other fields besides 3D room interiors; therefore it is not limited to this domain. Scientists may look at case studies in fashion design, product origination, or architectural visualization, among others, by refining fine techniques to meet the uniqueness of those datasets and tasks.

5. Collaborative Design Platforms:

Platforms of this sort that aggregate and filter design information through the use of generative models could be a trend in the future which would improve the interaction between designers and customers. Stakeholders could give real-time feedback, and also their opinions could be processed rapidly, so the final designs would be more effective and satisfying.

IX. CONCLUSION

The approach provided here is the one, called Stable Diffusion and Dreambooth, which is used to bring up the room inside generation process to the new level. Through the adaptation of Stable Diffusion with Dreambooth’s fast image generation capacity we created a platform of tailored instruments for the designing industry, like virtual design, interior design and architecture. It is a solution that meets the growing demand across disciplines for fast and personalized development.

Room development was a labor-intensive and time-consuming process, which, to a certain degree, predetermined constraints on people's imaginations and scalability. In contrast to this, a new spatial relation has occurred which got initiated with the introduction of deep learning and generative models such as Stable Diffusion and Dreambooth. These tools through being powerful, customizable and complete empower creative designers and creators to play professionally without losing originality and to keep their production rates high while expanding.

Stable diffusion has mastered in delivering various and near-real-looking images by making use of diffusion models that, unlike others, solve the problem of mode collapse. By the same token, Dreambooth connects to Stable Diffusion in which end-users are enabled to get involved in influencing and altering the process of generating room interior designs that are client-focused and customer-specific.

By fusing the inherent features of Stable Diffusion and Dreambooth, our creation which is an intelligent technological solution to one of the toughest hurdles of room interior generation with high efficiency and customer specificity. This combination of technologies has the capability of taking over design processes, improving outputs and, which in turn, will increase satisfaction of the end user as result.

The rise of Augmented Reality, Virtual Reality, and 3D Technology has signified a major shift in our perception of room interior designing in the digital era. These technologies bear out the limitless prospect and we are looking to uncover ground-breaking changes from it. An endless plethora of ideas can be put to work to enhance the residential feel that these futuristic spaces can provide.

In essence, this is the ultimate merger of Stable Diffusion and Dream Booth that opened a new gate in the history of room interior imagery creation, which guarantees saving of the cost and time, and also stimulates the creative aspect of the process.

X. REFERENCES

1. Zhang, L., Rao, A., & Agrawala, M. (2023). Adding Conditional Control to Text-to-Image Diffusion Models. -. https://doi.org/10.1109/iccv51070.2023.00355
2. Budige, U., & Konda, S. G. (2023, October 31). Text To Image Generation By Using Stable Diffusion Model With Variational Autoencoder Decoder. International Journal for Research in Applied Science and Engineering Technology, 11(10), 514–519. https://doi.org/10.22214/ijraset.2023.56024
3. Ghosh, S., Razniewski, S., & Weikum, G. (2020, October). Uncovering hidden semantics of set information in knowledge bases using GAN. Journal of Web Semantics, 64, 100588. https://doi.org/10.1016/j.websem.2020.100588
4. Moudgil, S., Gupta, V., & Sadana, R. (2019). Image style transfer using convolutional neural networks based on transfer learning. International Journal of Computational Systems Engineering, 5(1), 53. https://doi.org/10.1504/ijcsyse.2019.10019689
5. Yang, S., & Zhou, Y. (2022). Arbitrary Style Transfer with Style Enhancement and Structure Retention. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4305461
6. Zhang, C., Xu, Y., & Shen, Y. (2021). Decorating Your Own Bedroom: Locally Controlling Image Generation with Generative Adversarial Networks. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2105.08222
7. Abdal, R., Qin, Y., & Wonka, P. (2019). Image2StyleGAN: How to embed images into the StyleGAN latent space? arXiv (Cornell University). https://doi.org/10.48550/arxiv.1904.03189
8. T. Karras, T. Aila, S. Laine, and J. Lehtinen, “Progressive Growing of GANs for Improved Quality, Stability, and Variation,” arXiv.org, 2017. https://arxiv.org/abs/1710.10196
9. I. Goodfellow et al., “Generative adversarial networks,” Communications of the ACM, vol. 63, no. 11, pp. 139–144, Oct. 2020, doi: https://doi.org/10.1145/3422622.

10. Nauata, N., Chang, K. H., Cheng, C. Y., Mori, G., & Furukawa, Y. (2020). House-GAN: Relational Generative Adversarial Networks for Graph-Constrained House Layout Generation. Computer Vision – ECCV 2020, 162–177. https://doi.org/10.1007/978-3-030-58452-8_1