Journal of Artificial Intelligence Practice (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/jaip.2024.070124
ISSN 2371-8412 Vol. 7 Num. 1

A Multimodal Diffusion-based Interior Design AI with ControlNet

Ziyi Qin[1,a,*,#]

[1]*Wuxi Jiangsu Tianyi High School, Wuxi, China*
[a]*curryno915@gmail.com*
[#]*Co-first author*
[*]*Corresponding author*

Keywords: ControlNet, Diffusion Model, Interior Design

Abstract: This study introduces an AI-driven method for generating interior design visualizations using ControlNet, which effectively interprets blueprints and sketches to produce high-quality visualizations. By focusing on improving PSNR and SSIM metrics, the approach ensures structural accuracy and aesthetic appeal, demonstrating ControlNet's capability to capture detailed design information. The results showcase the model's potential to assist designers in creating refined and visually coherent interior design concepts, highlighting the benefits of AI integration in the design process.

1. Introduction

In the contemporary discourse of interior design, the integration of Artificial Intelligence (AI) represents a paradigm shift that is redefining the boundaries of creative expression and technical proficiency. This paper aims to critically assess the intrinsic value and future prospects of AI-generated interior design, acknowledging the transformative impact it has on the design community and the broader implications for the built environment.

The introduction of AI into the design process has introduced a novel dimension of efficiency, enabling rapid generation of design alternatives and facilitating a more personalized approach to space planning and aesthetics. This technological evolution not only expedites the design process but also democratizes access to professional design services, making them more affordable and attainable for a wider audience.

Moreover, the sustainability aspect of AI in interior design cannot be overlooked. By optimizing the use of space and materials, AI-driven designs contribute to the creation of eco-friendly and resource-efficient interiors, aligning with the global movement towards environmental consciousness in architecture and design.

Central to this exploration is the concept of ControlNet, a framework that allows designers to exert a level of oversight and refinement over AI-generated designs. This mechanism ensures that the output aligns with the designer's vision, respects practical constraints, and caters to the specific needs and preferences of clients, thereby maintaining the human-centric ethos at the core of interior design practice.

This paper will dissect the complex interplay between AI and human creativity, examining the ways in which ControlNet acts as a conduit for designers to harness the power of AI while preserving the integrity and artistry of their work. Through this lens, we will explore the trajectory of AI in interior design, charting its course from a nascent technology to a pivotal tool that shapes the aesthetic and functional fabric of our interiors.

2. Literature Review

In the landscape of network control systems, traditional centralized approaches face inefficiencies due to manual configurations. Recent studies have explored advanced techniques to automate the design process of network architectures [1,2]. One such approach, proposed by Wang et al., introduces time-aware DAMA-SGI, a method for modeling and analyzing irregularly sampled multivariate time series [3]. This method employs a joint task learning architecture, incorporating intra-sequence attention modules and learnable encoding of continuous time values to address asynchronous problems.

Neural Architecture Search (NAS) methods traditionally start with manually designed search spaces, limiting the final model's performance. However, Xiang et al. propose a graph-normalized flow-based generation method that leverages information from known high-performance models to generate architectures beyond predefined search spaces [4].

Introducing ControlNet, a decentralized system integrating machine learning methodologies, promises scalability and adaptability in network management [5]. By learning from network patterns and preemptively addressing issues, ControlNet aims to minimize manual intervention and optimize network performance.

In the domain of interior design, automated design processes hold promise for enhancing creativity and efficiency. Techniques such as Generative Adversarial Networks (GANs) have been explored for generating interior design layouts and furniture arrangements [6]. Additionally, StyleGAN, a GAN variant, has shown potential in synthesizing high-resolution and realistic images, which could be adapted for interior design applications [7].

The development of complex deep learning models has enabled new ways to approach the problems. For example, Zhou et al. utilized a graph-based deep neural network for hierarchical prediction of pedestrian trajectories [8]. Recent advancements in neural architecture search include differentiable architecture search (DARTS), which optimizes neural architectures directly on the task objective, eliminating the need for hand-crafted search spaces [9]. Reinforcement learning-based methods, such as Proximal Policy Optimization (PPO), have also been employed to efficiently explore vast architecture spaces [10].

Addressing the need for adaptable network control systems, approaches like Reinforcement Learning for Network Control (RLNC) have emerged, enabling networks to autonomously optimize performance and adapt to dynamic environments [11]. Furthermore, software-defined networking (SDN) coupled with reinforcement learning techniques offers a promising avenue for enhancing network flexibility and efficiency [12].

Generative models such as Variational Autoencoders (VAEs) have been applied to create novel interior designs by learning latent representations of furniture and room configurations [12]. Attention mechanisms, such as self-attention and Transformer architectures, have shown efficacy in generating coherent and context-aware interior layouts [13].

By integrating these advanced techniques, ControlNet can potentially revolutionize interior design by automating the creation of aesthetically pleasing and functional spaces while minimizing manual intervention and maximizing adaptability to evolving design preferences and spatial constraints.

3. Method

3.1. ControlNet

ControlNet is a neural network architecture which controls diffusion models by adding extra conditions. It can be used to generate images with greater precision and control by allowing users to add conditions such as canny edges and human poses.

The architecture of ControlNet consists of two main components: a diffusion model and a control network. The diffusion model is a pre-trained generative model that generates images from noise. The control network is a neural network that takes in an input image and a prompt and produces a synthesized image that matches the prompt and follows the constraints imposed by the input image.

The control network is trained using a combination of adversarial loss, perceptual loss, and feature matching loss. The adversarial loss ensures that the synthesized image is realistic, while the perceptual loss ensures that the synthesized image matches the prompt. The feature matching loss ensures that the synthesized image follows the constraints imposed by the input image.

3.2. Evaluation

In this part, we use two metrices to evaluate the validation result through the training process: The Structure Similarity (SSIM) as Eq.1 shows, and the Peak to Noise Ratio (PSNR) as Eq.2 shows.

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{1}$$

$$\mathrm{PSNR}(x,y) = 10\log_{10}\left(\frac{255^2}{\mathrm{MSE}(x,y)}\right) \tag{2}$$

4. Results

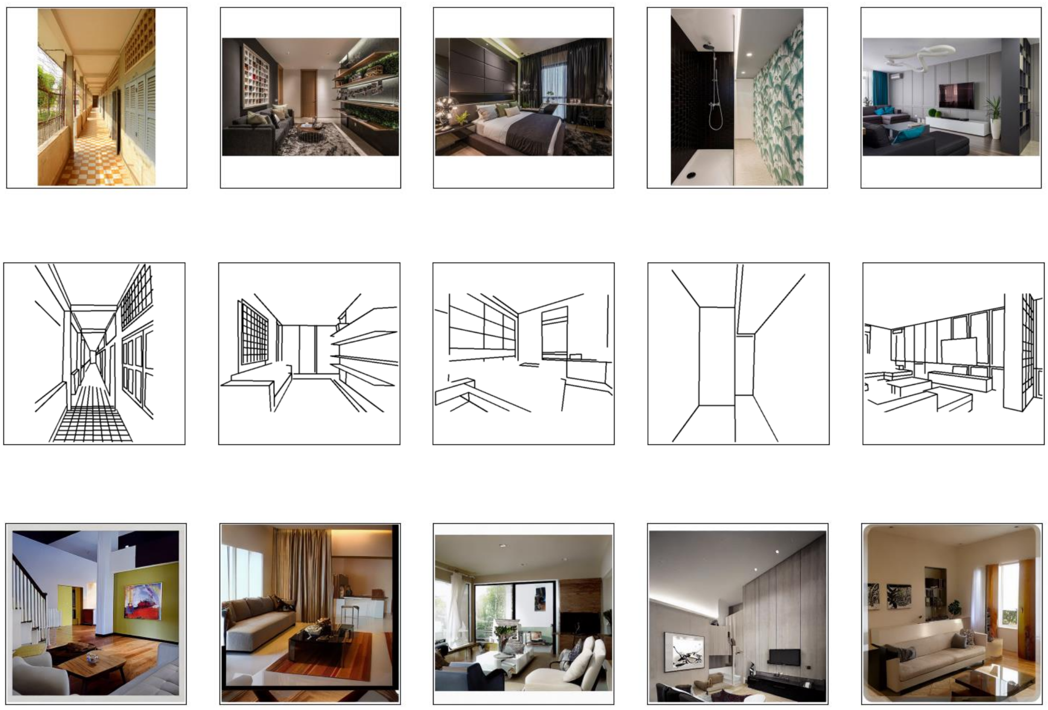

Figure 1: The ground truth, condition and generated image

Finally, the graphs we generated using ControlNet are comparable to the real graphs, the first row in Figure 1 is the ground truth and the third column is the image we generated using ControlNet. The PSNR of the generated image is 5.55 and the SSIM value of our images is 0.03. This result is reasonable, since the distribution of interior design images is far more complex than that of canny wireframes, which means two interior design images with utterly different colors and furniture elements may have similar wireframes. Since the prompt we use is merely "Interior design", the generation is highly variable. However, as shown in Figure 1, ControlNet does capture the structural information, as the generated images align with their ground truth in the general direction and layout.

Since the achieved PSNR and SSIM values are not so promising, there is still room for improvement. Future research will focus on optimizing the control network architecture to enhance these metrics.

5. Conclusion

ControlNet's innovative approach to network management, combining machine learning algorithms with a decentralized architecture, has proven effective in enhancing network performance and security. The framework's multi-layered design offers a comprehensive solution for real-time network resource optimization. Experimental results indicate significant improvements in network efficiency and resilience, suggesting ControlNet as a strong candidate for future smart network systems. Despite the need for further optimization, ControlNet sets a new standard for network control mechanisms, paving the way for more advanced and adaptive network management solutions.

References

[1] Huang, H., Amor, B.B., Lin, X., Zhu, F., Fang, Y. G-VAE, a Geometric Convolutional VAE for ProteinStructure Generation, arXiv.org, https://arxiv.org/abs/2106.11920, 2021.

[2] Salehin, I., Islam, M.S., Saha, P., Noman, S.M., Tuni, A., Hasan, Md.M, Baten, M.A., AutoML: A Systematic Review on Automated Machine Learning with Neural Architecture Search, Journal of Information and Intelligence, 2023.

[3] Wang, Z., et al., Toward Learning Joint Inference Tasks for IASS-MTS Using Dual Attention Memory with Stochastic Generative Imputation. IEEE Transactions on Neural Networks and Learning Systems, 2023.

[4] Yin, W., Yin, H., Kragic, D., Björkman, M., Graph-based normalizing flow for human motion generation and reconstruction, arXiv.org, https://arxiv.org/abs/2104.03020, 2021.

[5] Lvmin Zhang and Anyi Rao and Maneesh Agrawala. Adding Conditional Control to Text-to-Image Diffusion Models. International Conference on Computer Vision (ICCV), 2023.

[6] Dang, Q., et al., Towards an automatic interior design system using GAN. 170-172.10.1145/3478905.3478939. 2021.

[7] Karras, T., Laine, S., Aila, T., A Style-Based generator architecture for generative adversarial networks, arXiv.org, https://arxiv.org/abs/1812.04948, 2018.

[8] R. Zhou, H. Zhou, H. Gao, M. Tomizuka, J. Li, and Z. Xu, "Grouptron: Dynamic multi-scale graph convolutional networks for group-aware dense crowd trajectory forecasting," in 2022 International Conference on Robotics and Automation (ICRA), 2022, pp. 805–811.

[9] Liu, H., Simonyan, K., Yang, Y., DARTS: Differentiable Architecture Search, arXiv.org, https://arxiv.org/ abs/ 1806. 09055, 2018.

[10] Zoph, B., Le, Q., V., Neural Architecture Search with Reinforcement Learning, arXiv.org, https://arxiv.org/ abs/1611.01578, 2016.

[11] Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D., Continuous control with deep reinforcement learning, arXiv.org, https://arxiv.org/abs/1509.02971, 2015.

[12] Stampa, G., Arias, M., Sanchez-Charles, D., Muntes-Mulero, V., Cabellos, A., A Deep-Reinforcement learning approach for Software-Defined Networking routing optimization, arXiv.org, https://arxiv.org/abs/1709.07080, 2017.

[13] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017, June 12). Attention is all you need. arXiv.org. https://arxiv.org/abs/1706.03762