A FRAMEWORK FOR FINE-TUNING URBAN GANS USING DESIGN DECISION DATA GENERATED BY ARCHITECTS THROUGH GANS APPLICATIONS

XIMING ZHONG[1], JIADONG LIANG[2], FRICKER PIA[3] and SHENGYU LIU[4]
[13]Aalto university school of art design and architecture.
[2]Architectural Association School of Architecture.
[4]Politecnico di milano.
[1]ximing.zhong@aalto.fi, 0000-0002-3060-6981
[2]liangjiadongarch@gmail.com, 0009-0005-9244-8674
[3]pia.fricker@aalto.fi, 0000-0001-9748-6841
[4]shengyu.liu@mail.polimi.it, 0000-0002-0177-2948

Abstract. Recent studies have utilized Generative Adversarial Networks (GANs) to learn from existing urban layouts for urban design tasks. We define these GANs as Urban-GAN. However, urban layouts generated by Urban-GAN lack specificity and often require multiple modifications by architects to meet specific design requirements, making the process inefficient and non-customizable. Inspired by the concept of fine-tuning language models, we propose a stacked GAN model framework that fine-tunes Urban-GAN using data generated by architects in solving specific design tasks, forming AD-Urban-GAN. Our results indicate that layouts produced by AD-Urban-GAN more effectively emulate architects' design morphology decisions, enhancing Urban-GAN's adaptability and efficiency in handling design tasks. Furthermore, AD-Urban-GAN enhances the customizability of Urban-GAN models for specific urban design tasks, generating layouts that accurately understand and meet the requirements of specific tasks. AD-Urban-GAN significantly streamlines the process of generating design prototypes for specific task types, enabling precise quantitative control over urban layout results. This workflow establishes a data acquisition and training loop that strengthens the customizability of existing GANs. The design decision data generated by architects can improve the adaptability and customization of GANs models, facilitating efficient collaborative work between architects and artificial intelligence.

Keywords. Architect Design Decisions, Fine-tuning, GANs, Stack-GANs, Adaptability, Customizability

1. Introduction

Generative Adversarial Networks (GANs) is a generative model that consists of a discriminator and a generator, with the generator generating realistic data and the

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 2, 19-28.

discriminator distinguishing as accurately as possible between generated and real data (Goodfellow et al. 2020). In architectural design for plan layout and elevation design, GANs become an important generative tool (Chaillou,2020). Recent studies have leveraged GANs to learn from existing urban layout schemes for addressing urban design tasks (Shen et al., 2020; Tian, 2021). We define these GANs as Urban-GAN. Tools like Urban-GANs, while capable of simulating urban layouts similar to historical datasets, include conditional GANs that allow architects to input site constraints and interactively control the generation of urban layouts. Despite this, architects still require extensive interaction and exploration to finalize plans based on complex site constraints and implicit conditions (Li et al., 2020).

Urban layouts generated by Urban-GAN lack specificity and often require multiple revisions by architects to meet specific design needs, a process that is both inefficient and non-customizable. For instance, Zhong et al. (2022) implementation of Urban-GAN still exhibits considerable randomness in the generated forms, necessitating the input of numerous conditional constraints by architects to control GAN outcomes and ensure the rationality and customization of design requirements. This is because Urban-GAN models often lack a clear mechanism for deeply understanding and adapting to complex design sites. In contrast, architects, when addressing particular design tasks, are capable of extracting key features from on-site information, thereby creating solutions that are not only highly adaptive but also personalized and rational in form. In the realm of urban design, the decision-making process employed by architects in generating design solutions surpasses the current capabilities of machine-generated decisions. (Davies et al., 2021).

The concept of fine-tuning has recently been extensively employed to train general-purpose language models for specific tasks, which, compared to general models, enables more efficient and precise solutions to certain types of problems (Sun et al., 2019). Inspired by this concept of fine-tuning, and seeking to combine the advantages of artificial intelligence with human decision-making (Davies et al., 2021), we aim to utilize data from the interactions between architects and GANs for training AI. This approach is intended to enhance the capability of Urban-GAN to generate layouts that simulate human architect design decisions, thereby obtaining a more customized and adaptive GAN model.

Stack-GAN is a framework that employs a stacked approach to GANs. It comprises two GANs: the first captures basic features of the data, whose output is then fine-tuned by the second GAN, using it as its input (Zhang et al., 2017). This inspires us to train a GAN using actual data generated by architects while handling specific design tasks(AD-GAN), and then stack it with the Urban-GAN model. This combination forms a GAN that not only learns historical layout features but also simulates rational morphology control by human architects(AD-Urban-GAN), as shown in Figure 1. To integrate multiple architects' design decisions into a single design task, we also employ Principal Component Analysis (PCA), a data dimensionality reduction and feature extraction technique. PCA represents data with fewer dimensions while retaining as much of the original data's information as possible, integrating common features from multiple datasets (Wold, S, 1987). This helps us extract important common features from a vast array of architects' design decision characteristics, further enhancing the adaptability and customization capabilities of AD-Urban-GAN.

A FRAMEWORK FOR FINE-TUNING URBAN GANS USING DESIGN DECISION DATA GENERATED BY ARCHITECTS THROUGH GANS APPLICATIONS

We propose the AD-Urban-GAN framework, a combination of AD-GAN and Urban-GAN. Compared to Urban-GAN, AD-Urban-GAN can more closely emulate the architect design morphology decisions, reducing the frequency of modifications required by architects on Urban-GAN-generated results, and thereby increasing the efficiency of architects in handling design tasks. Additionally, the layout forms produced by AD-Urban-GAN are more controllable and closely align with the quantitative design metrics of specific design tasks. This enhances the customization capabilities of the URBAN-GAN model, enabling the generated layouts to more accurately understand and respond to the design requirements of specific tasks.

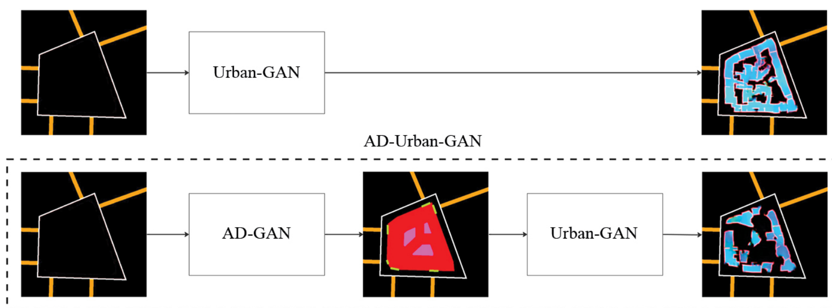

Figure 1. User Interface for Urban Layouts Generation Using Employed AD-Urban-GAN

2. Related Work

Tian (2021)employs conditional GANs to learn the multiple constraints of a site from historical urban layout data, allowing architects to generate layouts by controlling labels. However, due to the limitations of the label information, precise control over the architectural layout within boundaries is unachievable. Zhong et al.(2022) enhance the control over GAN-generated layouts by using hierarchical urban features as data labels, allowing architects to interactively control labels and customize the generation of three-dimensional urban layouts. Nonetheless, these data labels, derived from past historical case data, make it challenging to meet the current site's constraints with label-controlled results. Architects still need to understand the site conditions and continuously input labels to control GANs. We aim for GANs to truly simulate architects' design decisions, enhancing the design efficiency and customization capabilities of the GAN model.

Stack-GAN generates high-resolution, realistic images from textual descriptions, whereas the first GAN produces low-resolution images with basic shapes and colours, and the second GAN utilizes these low-resolution images and text descriptions to generate high-resolution images with realistic details (Zhang et al., 2017). Stack-GAN establishes a hierarchical structure between the two stages of generators and discriminators, using conditional augmentation techniques to ensure that the final generated images meet the design requirements. This method inspired us to stack AD-GAN and Urban-GAN models, inserting architectural design decision constraints into the hierarchical structure to fine-tune urban layout generation results. Concurrently, we use PCA to extract common features of multiple architects' design decisions to train

ADGAN, better simulating architects' design strategies. Härkönen et al. (2020) use PCA to identify key attribute variations in facial feature components, demonstrating its effectiveness in identifying important features in data. Additionally, Koh (2019) utilized PCA to extract spatial layout features of multiple churches, and then generate new floor plans by piecing together various church space features in different proportions. These examples demonstrate the potential of PCA in extracting and reassembling complex design features, particularly showcasing its robust capability when dealing with diversified and high-dimensional data, allowing us to process large datasets from architects for feature mining and labeling.

The primary experimental objectives of this paper are: To demonstrate that the AD-GAN, trained with actual architectural design decision data and stacked with Urban-GAN, can simulate architects' design decisions for specific design tasks. To show that layouts generated by AD-Urban-GAN are more precise and controllable, meeting the design metrics of specific design tasks.

3. AD-URBAN-GAN Framework

The AD-Urban-GAN framework is illustrated in Figure 2.

Figure 2. AD-Urban-GAN Workflow

3.1. URBAN-GAN FRAMEWORK

We adopted Zhong's GAN model as the foundational Urban-GAN model for our research, specifically, the input was morphological colour coding (MCC) and the output was a 3D volume with different functions (Zhong et al. 2022). The MCC contained RGB (Fig 3). We obtained 4000 pairs of image data from OSM, which

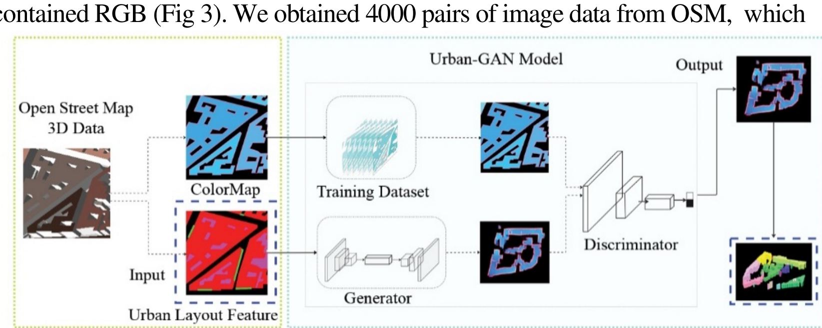

Figure 3. Urban-GAN for Learning Historical Case Data

included site boundaries, building heights, and functional classification information. The results of Urban-GAN were transferred to 3D models by decoding the MCC map. The accuracy of this approach has been demonstrated in previous work (Tian, 2021). The architect used different inputs of morphological decision colour maps to complete the simulation of the layout task.

3.2. ARCHITECT DESIGN DECISIONS DATA FOR AD-GAN TRAINING

To collect a dataset based on customized design tasks, we asked architects to draw design decision diagrams online, as shown in Figure 4(b)(c). Architects could refer to layouts generated by Urban-GAN to design forms that met the morphological metrics required by different tasks, as illustrated in Figures 4(a)(d)(e). We recorded the morphological metrics of the design decision diagrams learned by Urban-GAN, the morphological metrics annotated by architects, and the morphological metrics that needed to be met by the design task.

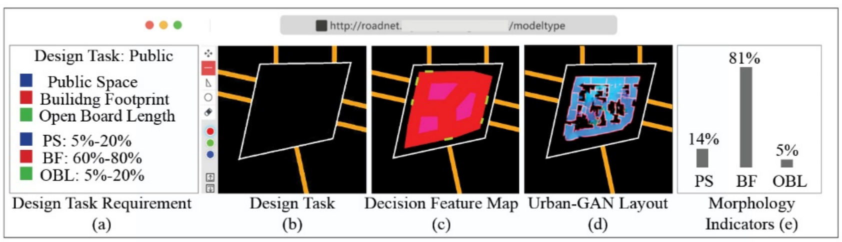

Figure 4. Architect Design Decisions Data for Specific Tasks

Figure 5 shows architects' design layouts under all task types from 85 professional architects and students from an architecture university in China, with a total of 1800 valid solutions obtained and used to train the AD-GAN model, and 480 cases used for the test dataset.

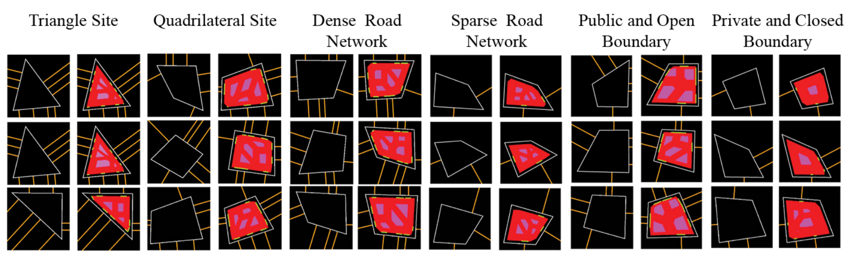

Figure 5. Dataset for Different Tasks

3.3. PCA EXTRACTING FEATURE VECTORS FROM DESIGN DECISIONS

Figure 6 demonstrates the process of using the PCA to integrate the common features of the design intuitions of different architects for the same site. For each task, we processed the data as a batch with PCA to reduce noise and prevent overfitting

(Moschoglou et al. 2020). Experiments showed that the feature vector was compressed to 10 to minimize the loss of data while minimizing the dimensionality of the data. The PCA extracted the important features of multiple architects' design strategies for the same community design. If we needed to modify the results of AD-GAN, the 10 principal feature vectors obtained from the PCA calculation (Fig.6(c)) could be formed by linearly stacking them with different weights. We continuously stacked the features by adjusting the feature parameters until we obtained the AD-GAN morphological decision that met the design requirements.

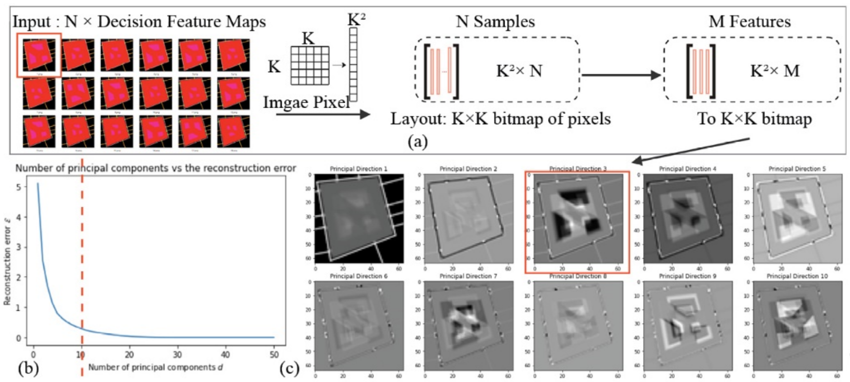

Figure 6. Data Processing by PCA (a) Compression of Design Decisions Data (b) Set Feature Vectors According to Reconstruct Errors Parameters (c) Compression of Vector Features

3.4. STACKING AD-GAN AND URBAN-GAN

Inputting design tasks, we got the feature map by the trained AD-GAN (Fig.7(a)), and then PCA integrated the feature map and decision data of multiple architects to generate a new layout with different weight proportions. The layout combined with ten feature vectors extracted from the collective intuition decision data to adapt to the changes in decision-making during the design process, passing down to Urban-GAN(Fig.7(b)).

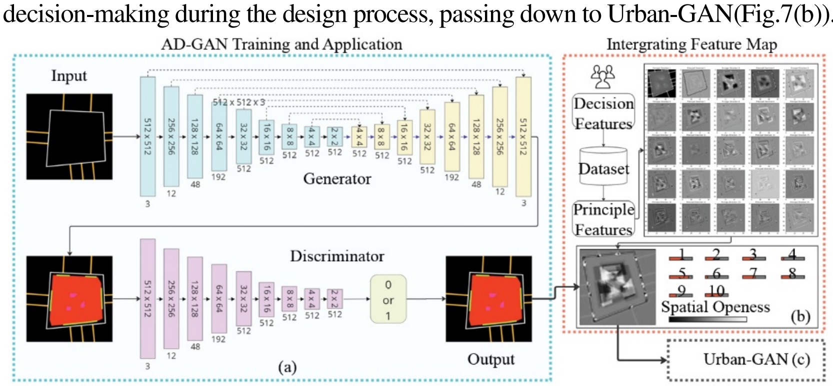

Figure 7. AD-GAN Integrated with Architect Design Decisions Stacking Urban-GAN

4. Result

4.1. RESULTS OF AD-GAN GENERATION AND DESIGN DECISIONS

The experimental results demonstrated that the layouts generated by AD-GAN could simulate architects' design decisions. As illustrated in Figure 8, facing different intersections and form boundaries, the results of AD-GAN and the actual results exhibited remarkable similarity. Specifically, the bar graph (Fig.8(4)) showed the comparison of morphological metrics for different types of tasks between AD-GAN and architects. The morphological metrics of layouts generated by AD-GAN, including BF, PS and OBL, were highly similar to those in architects' design decisions, with an average difference in data comparison ranging between 0.025 and 0.041.

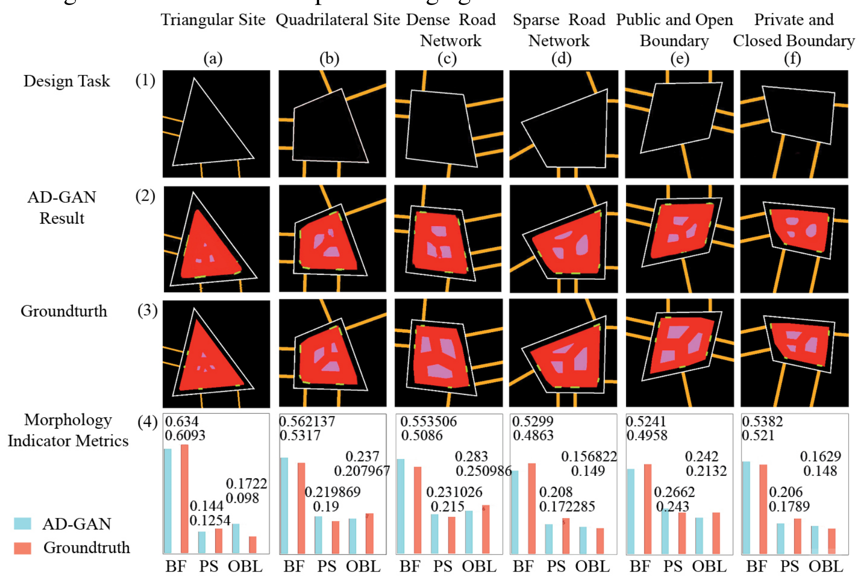

Figure 8. AD-GAN and Ground Truth Test Results

4.2. COMPARISON RESULTS OF AD-URBAN-GAN AND URBAN-GAN

By comparing the output results and morphological metrics of AD-Urban-GAN and Urban-GAN in urban tasks, we observed that AD-Urban-GAN's layouts were more rational and met task metrics. Figures 9 (1) and (2) revealed that AD-Urban-GAN, effectively reduced irregular internal spaces, resulting in more orderly and reasonable divisions of public spaces. In terms of coordination with road networks, layouts generated by AD-Urban-GAN showed superior performance, indicating that AD-Urban-GAN could more deeply simulate architects' handling of the complex relationship between roads and open building boundaries. Morphological metrics analysis, as shown in Figure 9(5), indicated that the morphological metrics of layouts generated by AD-Urban-Gan were closer to the design task requirements. For example, in public space-type site tasks (Fig.9 (e-5)), the public space (PS) value of layouts generated by AD-Urban-Gan reached 0.195, significantly surpassing the PS value of Urban-GAN layouts and meeting the design task requirement of 0.18-0.2 PS. Moreover, in densely networked plots, AD-Urban-Gan's results had an advantage in terms of open boundary proportion (Fig.9 (c-5)), with an Open Boundary Line (OBL) value of 0.169, which was 0.11 higher than that of Urban-GAN and close to the task requirement of 0.15-0.18 OBL. Figure 9 (3) showed a comparison between the layouts generated by AD-Urban-Gan and the marked diagrams generated by AD-GAN simulating architects' design decisions, clearly demonstrating AD-Urban-Gan's simulation of architects' design strategies and their alignment with site information.

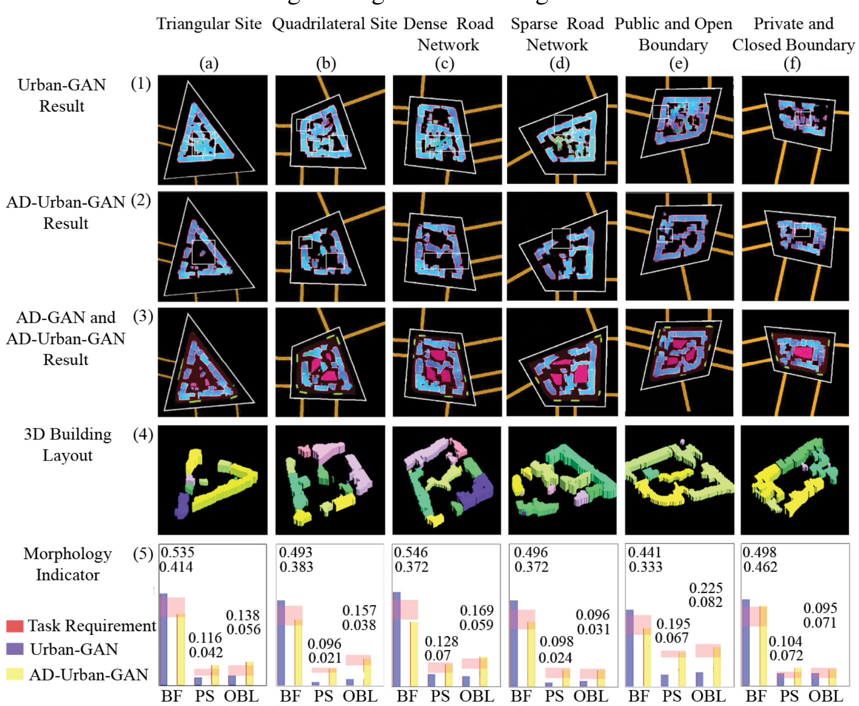

Figure 9. AD-Urban-GAN and Urban-GAN Test Results

4.3. FINE-TUNING AD-URBAN-GAN GENERATED LAYOUT RESULTS

The experimental results indicated that PCA combined the design decision features of multiple architects, finely tuning the results generated by AD-Urban-GAN through detailed parameters. PCA was employed to linearly combine architectural design decisions from 30 architects with varying weight proportions, iteratively generating new design layouts(Fig.10). We demonstrated the changes in layout from different proportional combinations of feature vectors 5, 7, and 10. Feature vector 5 represented public spaces with lower openness and dispersed independence, feature vector 7 represented highly open, clustered public spaces and feature vector 10 represented the rotation of public spaces. The combination of features 5 and 10 showed a clockwise change in the layout of public spaces while maintaining their independent arrangement. The combination of features 5 and 7 illustrated a gradual transition from the independent public space layout of feature 5 to the clustered layout of feature 7. We utilized PCA to extract multiple architects' design decision features for the same design task. These features were combined and controlled the direction and variation of AD-GAN results, better fine-tuning the layouts generated by AD-Urban-GAN.

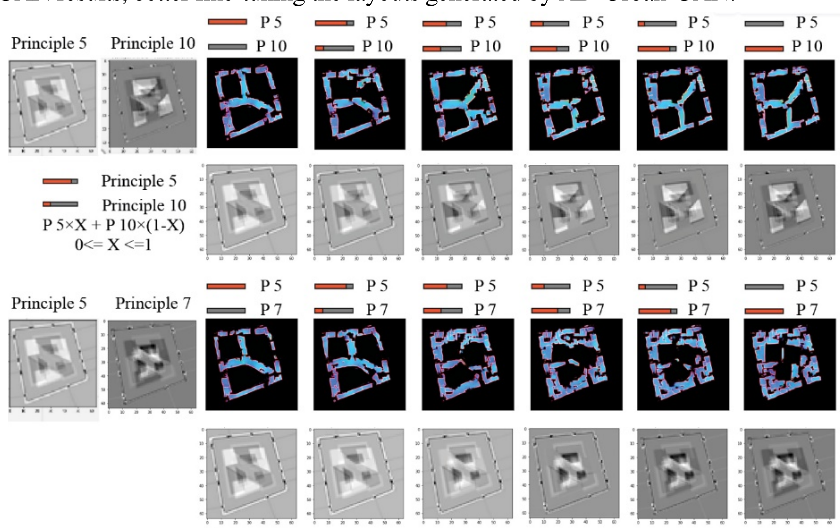

Figure 10. Fine-Tuning of AD-Urban-GAN for Generating New Layout Results

4.4. LIMITATION

GANs currently lack interpretability in understanding the complexity of sites. Graph Neural Networks (GNNs) have shown the ability to interpret site conditions (Xu et al., 2018). In future work, we plan to incorporate GNNs to learn the interrelations between site and conditions, combined with the generative capabilities of GANs to produce designs that meet practical task requirements. In our experiments, we simplified the design process using a limited dataset, enabling architects to control the GANs results. However, in real urban design tasks, more complex site conditions must be considered. In subsequent tasks, we will use more comprehensive environmental data to enhance the accuracy and design diversity of GAN-generated solutions in practical applications. Currently, GANs can only understand and address specific tasks determined by datasets. To extend the applicability of GANs to a broader range of design tasks, diverse data can be obtained through collaboration with architects, thereby facilitating improvements in model training and the promotion of practical applications.

5. Conclusion

This study introduces AD-Urban-GAN, a framework that stacks AD-GAN and Urban-GAN models, combining human architects' design decisions with existing urban layout data. It demonstrates formidable adaptability and customization capabilities in handling urban design tasks. Architects can significantly streamline the process of generating design prototypes based on specific task types using AD-Urban-GAN, achieving precise quantitative control over layout results. The application of AD-Urban-GAN facilitates a pivotal shift, enabling the formation of a comprehensive database containing design decision data during the interaction process between architects and machines. This workflow establishes a cycle of data acquisition and training to enhance the urban design capabilities of existing GANs.

References

Chaillou, S. (2020, September). Archigan: Artificial intelligence x architecture. In Architectural Intelligence: Selected Papers from the 1st International Conference on Computational Design and Robotic Fabrication (CDRF 2019) (pp. 117-127). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-15-6568-7_8

Davies, A., Veličković, P., Buesing, L., Blackwell, S., Zheng, D., Tomašev, N., Tanburn, R., Battaglia, P., Blundell, C., Juhász, A., Lackenby, M., Williamson, G., Hassabis, D., & Kohli, P. (2021). Advancing mathematics by guiding human intuition with AI. Nature, 600(7887), 70–74. https://doi.org/10.1038/s41586-021-04086-x

Härkönen, E., Hertzmann, A., Lehtinen, J., & Paris, S. (2020). Ganspace: Discovering interpretable gan controls. Advances in neural information processing systems, 33, 9841-9850. arXiv. http://arxiv.org/abs/2004.02546

Koh, I. C. B. (2019). Architectural sampling: a formal basis for machine-learnable architecture (No. THESIS). EPFL. https://doi.org/10.5075/EPFL-THESIS-7815

Li, M., Lin, J., Ding, Y., Liu, Z., Zhu, J.-Y., & Han, S. (2020). GAN Compression: Efficient Architectures for Interactive Conditional GANs. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 5283–5293. https://doi.org/10.1109/CVPR42600.2020.00533

Moschoglou, S., Ploumpis, S., Nicolaou, M. A., Papaioannou, A., & Zafeiriou, S. (2020). 3DFaceGAN: Adversarial Nets for 3D Face Representation, Generation, and Translation. International Journal of Computer Vision, 128(10–11), 2534–2551.

Shen, J., Liu, C., Ren, Y., & Zheng, H. (2020, August). Machine learning assisted urban filling. In Proceedings of the 25th CAADRIA Conference, Bangkok, Thailand (pp. 5-6).

Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to fine-tune bert for text classification?. In Chinese Computational Linguistics: 18th China National Conference, CCL 2019, Kunming, China, October 18–20, 2019, Proceedings 18 (pp. 194-206). Springer International Publishing. https://doi.org/10.48550/ARXIV.1905.05583

Tian, R. (2021). Suggestive site planning with conditional gan and urban gis data. In Proceedings of the 2020 DigitalFUTURES: The 2nd International Conference on Computational Design and Robotic Fabrication (CDRF 2020) (pp. 103-113). Springer Singapore. https://doi.org/10.1007/978-981-33-4400-6_10

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. Chemometrics and intelligent laboratory systems, 2(1-3), 37-52.

Xu, K., Hu, W., Leskovec, J., & Jegelka, S. (2018). How Powerful are Graph Neural Networks? https://doi.org/10.48550/ARXIV.1810.00826

Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., & Metaxas, D. N. (2017). Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks. In Proceedings of the IEEE International Conference on computer vision (pp. 5907-5915). arXiv. http://arxiv.org/abs/1612.03242

Zhong, X., Fricker, P., Yu, F., Tan, C., & Pan, Y. (2022). A Discussion on an Urban Layout Workflow Utilizing Generative Adversarial Network (GAN)—With a focus on automatized labeling and dataset acquisition.