https://doi.org/10.15302/J-LAF-1-020094 Received Date | 2023-11-26 Accepted Date | 2024-02-06

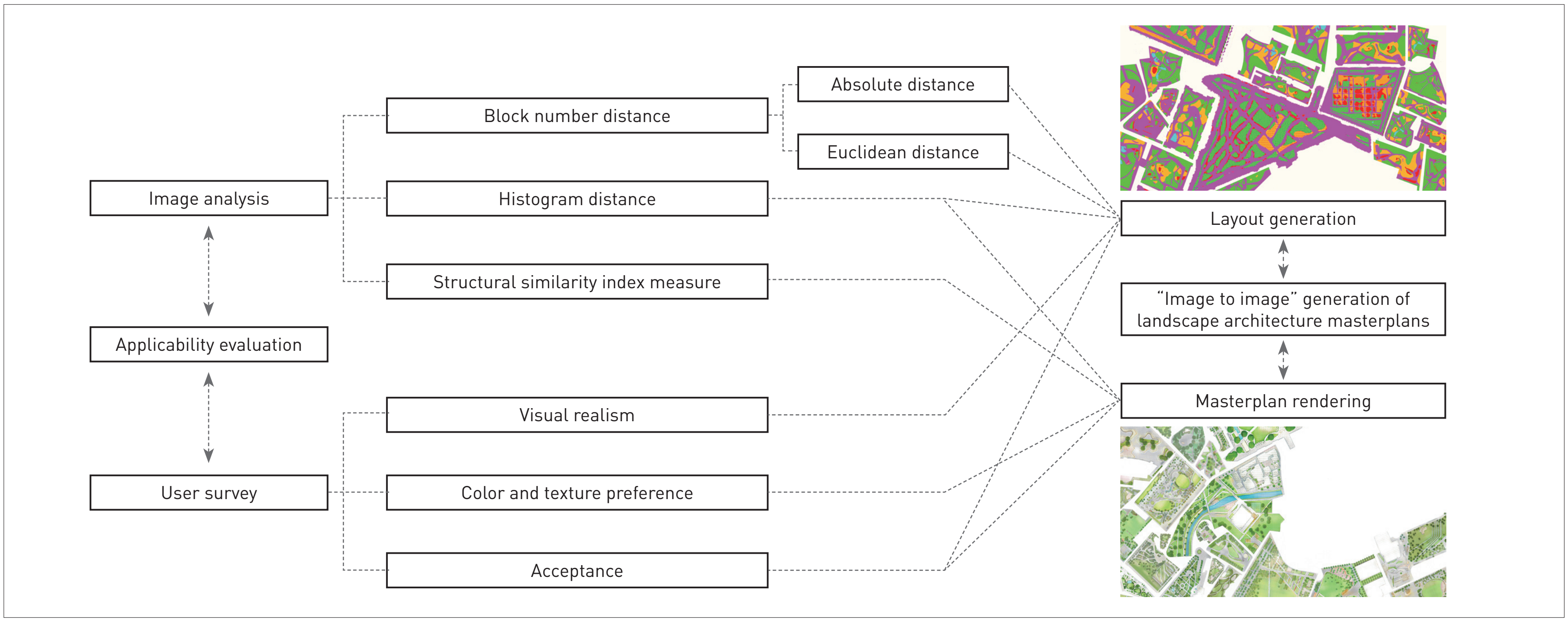
Applicability Evaluation and Reflection on Artificial Intelligence-based "Image to Image" Generation of Landscape Architecture Masterplans

Huaiyu ZHOU[1,2,*], Shuangbin XIANG[2,3]

1 Department of Architecture, School of Architecture and Planning, Hunan University, Changsha 410082, China
2 Department of Landscape Architecture, School of Architecture, Tsinghua University, Beijing 100084, China
3 Landscape Architecture Studio, Beijing General Municipal Engineering Design & Research Institute Co., Ltd., Beijing 100088, China

*CORRESPONDING AUTHOR
Address: School of Architecture and Planning, Hunan University, Changsha 410082, Hunan Province, China
Email: zhouhy@hnu.edu.cn

GRAPHICAL ABSTRACT



ABSTRACT

Artificial intelligence (AI) image generation is revolutionizing traditional workflow in landscape architecture industry, among which the "image-to-image" generative adversarial network (GAN) exhibits potential to facilitate concept design. Therefore, it underscores the importance of applicability evaluation from the perspective of users. This research aims to evaluate the quality of the GAN-generated results, their effectiveness in integrating with design workflows, and the landscape architects' acceptance of the results through image analysis and user survey. The evaluation focuses on layout generation and masterplan rendering within the Pix2Pix–BicycleGAN workflow. The evaluation metrics of image analysis including block number absolute/Euclidean distance, histogram distance, and structural similarity index measure, were employed. Additionally, the online survey with two questionnaires was conducted to evaluate the visual realism and preference for color and texture of the GAN-generated results. The findings indicate that the GAN-generated layout exhibits a high similarity to the human-designed layout, and the GAN-rendered masterplans fulfill the criteria for concept design and garner positive user acceptance. Conclusively, this study delves into the intrinsic rationality of the GAN generation methods and limitations in professional ethics and data bias, reflecting on the gaps between current AI-assisted design methods and evidence-based design.

KEYWORDS

Landscape Architecture; Image Generation; Generative Adversarial Network; Artificial Intelligence-Assisted Design; Applicability Evaluation; Landscape Masterplan

HIGHLIGHTS

- Quantitative applicability evaluation of "image to image" landscape masterplan generation method
- Image analysis reveals a high similarity between GAN-generated and human-designed layouts
- User survey reveals a high visual realism and practitioners' high acceptance of GAN-rendered masterplans
- Identifies the intrinsic rationality of current GAN generation methods and the technical gaps between these methods and evidence-based design

EDITED BY Yuting GAO, Ying WANG
TRANSLATED BY Yuting GAO

1 Introduction

In recent years, the rapid development and enhancement of image generation technologies and mapping tools driven by generative artificial intelligence (AI) have significantly impacted the traditional landscape design industry[1]~[3]. Thus, it is pressing for landscape architects to delineate the relationship between image generation and landscape design and explore potential opportunities of practice and research. The applicability evaluation from the landscape architects, the users of image generation technologies, can assist in analyzing its potential impact, optimizing tool selection, and ultimately enhancing design efficiency. Presently, image generation technologies are primarily applied in masterplan generation and perspective rendering in the landscape design workflow.

Research on masterplan generation primarily focuses on "image-to-image" generative adversarial network (GAN). The application of these tools has developed from the generation of architectural floor plans[4]~[6] to generating building arrangements and massing relationships[7]~[10]. In recent years, relevant research also initiated in the field of landscape architecture on masterplan generation: Huaiyu Zhou et al. established a labeled masterplan dataset and adopted CycleGAN① for landscape masterplan recognition and rendering[11]; Guangbin Qu et al. generated functional layouts of landscape with CGAN② in residential areas that meet design specifications[12]; Ran Chen et al. utilized StyleGAN2③ to generate diverse design schemes, discovering that GAN models can recognize and extract high-dimensional abstract features of vegetation, water bodies, pavements, and road networks[13]; Guanjie Zhao explored automated design processes for small-scale landscapes by coupling Pix2Pix④ and Stable Diffusion models[14]; Weishi Zhou used Pix2Pix to generate masterplans for urban pocket parks[15]. Despite the in-depth discussions on training principles, datasets, and generation methods in existing studies, several issues remain: there lacks publicly accessible landscape masterplan datasets, which limits the diversity of training data; the scale of the generable masterplans is constrained and mainly suitable for small- and medium-sized green spaces; the targeted systematic quantitative evaluation of GAN-generated masterplans is insufficient, lacking user-friendly evaluation metrics; and there are limited user-side surveys, making it difficult to obtain usage evaluations.

Relevant research and applications of rendering focus on two major "text-to-image" tools: Midjourney and Stable Diffusion. The Midjourney model can generate highly refined and realistic human perspective views or bird's-eye views on web platforms using various prompts, making it user-friendly. In contrast, the open-source Stable

① CycleGAN (Cycle Generative Adversarial Network) enables unsupervised image-to-image translation between two different image domains without requiring paired training data, making it suitable for style transfer.
② CGAN (Conditional Generative Adversarial Network) introduces additional conditions or labels in the image generation process, ensuring that the generated results are influenced not only by random noise but also by specific conditions.
③ StyleGAN (Style Generative Adversarial Network) enhances the diversity of generated images through the training of style vectors, excelling in image detail and large-scale dataset processing.
④ Pix2Pix ("Image-to-Image" Generative Adversarial Network) generates corresponding output images based on input ones, such as converting black-and-white images to color or converting sketches to photos, excelling in image synthesis tasks.

Diffusion model can not only generate images through keywords but also offer "image-to-image" and "model-to-image" training functions, allowing designers to add constraints, thereby gaining more popularity. Currently, a Stable Diffusion-based workflow for architectural form conception and modeling has been established[16][17].

As technology iterates and advances, design workflows have evolved from hand-drawing to CAD drafting and then to parametric design with Grasshopper. Designers have facilitated new workflows by integrating emerging technologies in practice and actively conducting evaluations[18]. This study focuses on GAN-based landscape masterplan generation methods ("GAN generation methods" hereafter), comprehensively assessing their technical applicability from the perspective of landscape architects to provide references for tool selection. The core advantage of GAN generation methods in "image-to-image" tasks is their mapping efficiency, which reduces the time consumption on concept comparisons and repeated rendering[11]. Therefore, this study aims to evaluate the quality of the GAN-generated results, their effectiveness in integrating with design workflows, and the landscape architects' acceptance of the results through image analysis and user survey.

2 Evaluation Object: Pix2Pix–BicycleGAN Workflow

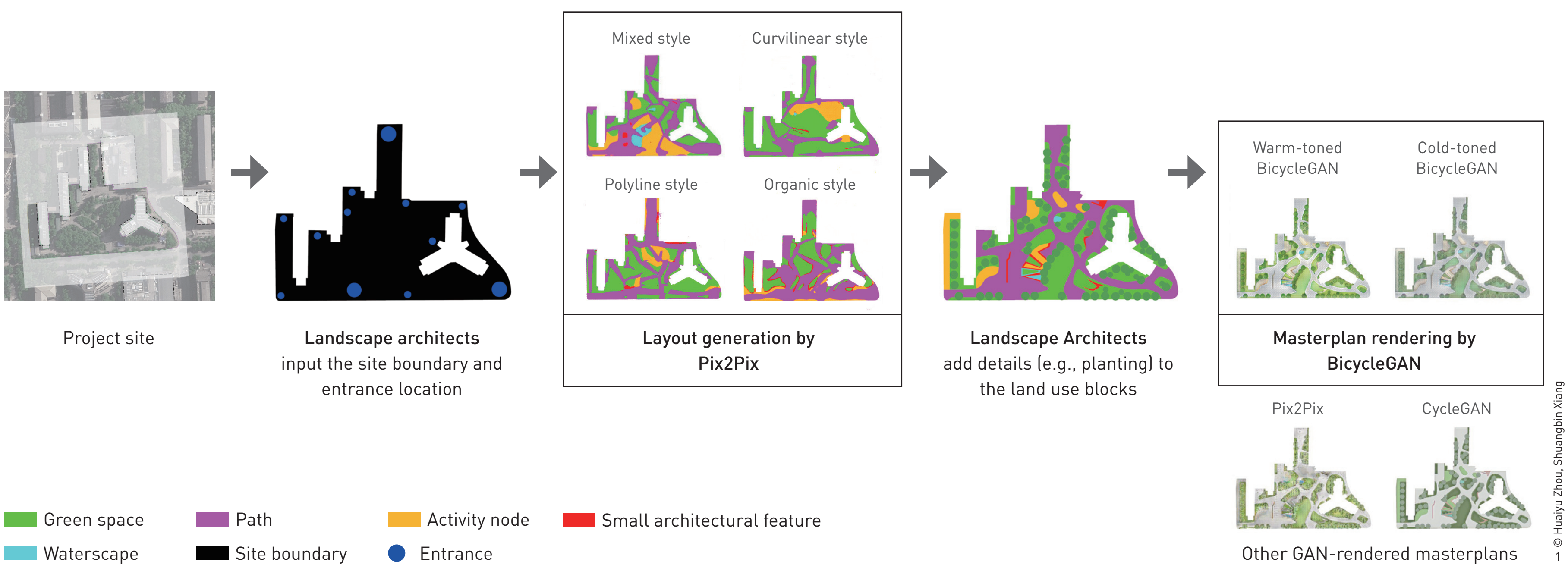
This study focuses on the adaptability evaluation of two key tasks in the Pix2Pix–BicycleGAN landscape masterplan generation workflow—layout generation and masterplan rendering. GAN-generated layouts are similar to functional bubbles and schematic sketches in design education, representing an intuitive and straightforward way of thinking that forms the basis for design iteration and adjustment. GAN-rendered images add details of color and texture to the abstract layout, enhancing their readability. Pix2Pix[19] is a widely used task implementation model in the GAN field; while BicycleGAN[20], an improved model of CycleGAN[21], introduces additional variables and constraints, improving the model's performance in handling multimodal data and high-resolution images and supports the output of various rendering results. Due to the limited types of masterplans in the collected and labeled dataset, this workflow is mainly applicable to small and medium-scale landscape presently[11][13]~[15] (Fig. 1).

2.1 Generating Layouts in Various Styles

By inputting the site boundary into the Pix2Pix model, layouts in various styles (mixed, curvilinear, polyline, and organic styles) containing different land-use types can be generated, including green spaces, paths, activity nodes, small architectural features, and waterscapes (Fig. 1). The evaluation of GAN-generated layouts focuses on their morphological similarity and visual realism compared to the human-designed layouts.

In this study, a total of 2,725 human-designed landscape masterplans were collected, with training sets of 2,670, 916, 770,

1. Example of layout generation and masterplan rendering in the Pix2Pix–BicycleGAN workflow.



and 954 images for mixed, curvilinear, polyline, and organic styles, respectively⑤. An additional validation set of 85 masterplans was reserved for evaluating the generation results. The converted site layouts in PNG or JPEG format display the site boundary filled in black with entrance locations marked with blue circles (their size indicating entrance level). Based on the four styles, 340 GAN-generated layouts (85×4) were gathered for subsequent evaluation. Landscape architects, after comparing multiple GAN-generated layouts, can continue to adjust the formal design, supplement planting, refine land use division, and form more refined site layouts based on project requirements and personal experience, using these as inputs for the masterplan rendering.

2.2 Rendering Masterplans with Various Colors and Textures

Landscape architects can input the adjusted layouts into BicycleGAN to generate rendered masterplans with different colors and textures[20], facilitating efficient communication of design concepts with clients. The evaluation of this task primarily focuses on the similarity between GAN-rendered masterplans and manually rendered masterplans (such as those colored by hand or using tools like Adobe Photoshop) and user preferences for colors and textures. The dataset included 325 landscape layouts, with a training set of 300 layouts and a validation set of 25 layouts[11]. Since BicycleGAN generates multiple rendering results, one warm-toned rendering and one cold-toned rendering were selected for each layout, totaling 50 outputs for evaluation. The rendered masterplans generated using the same layouts with Pix2Pix and CycleGAN models from the author's previous research[11] were also included in further image analysis and user survey⑥.

3 Evaluation Methods

The evaluation of the similarity between GAN-generated and human-designed layouts, as well as between GAN-rendered and manually rendered masterplans, focuses on the features of the images themselves and is therefore suitable for image analysis. Conversely, the evaluation of the visual realism of the generated layouts and preferences for the rendered masterplans is better suited to user surveys. Hence, this study integrates image analysis and user surveys to establish an evaluation metric system.

⑤ The Mixed training set is a collection of the curvilinear, polyline, and organic training sets.

⑥ For specific model training methods, see Ref. [11].

3.1 Image Analysis Metrics

3.1.1 Evaluation Metrics for Layout Generation

(1) Block number distance

The block number (BN) of the five generated land-use types can most directly reflect the morphological diversity of GAN-generated layouts. The corresponding block number distance (BND) can be used to assess the differences between the 340 validation set layouts generated by Pix2Pix and the human-designed layouts. BND evaluation includes the calculation of absolute BND and Euclidean distance. In this study, the absolute distance was used to compare the difference in the BN of various land-use types (manually counting complete color blocks with the same RGB value) between the generated layouts and the human-designed layouts for each single style. Additionally, to further analyze the impact caused by the different numbers of layouts of each style's training set, this study conducted cluster analysis of absolute BND and Euclidean BND to compare the differences in the clustering degree of land use division among the four styles. The midpoint clustering regions of the two sets of data were presented in a cluster analysis graph.

(2) Histogram distance

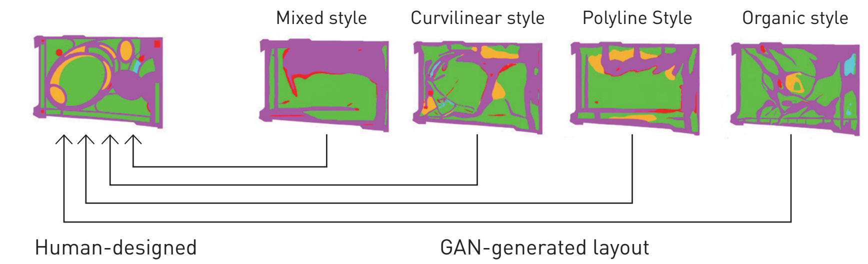
An image histogram shows the frequency distribution of different RGB pixels in an image. Histogram distance (*HistD*) is a key metric for measuring the pixel distribution differences between two images[22]. Based on the one-to-one correspondence between the RGB values in the GAN-generated layouts and the land-use types, *HistD* can effectively assess the variation in land use division and area proportion between GAN-generated layouts and human-designed layouts. In this study, Bhattacharyya distance was used to quantify the distance between two normalized histograms:

$$HistD = \sqrt{1 - \sum_i \sqrt{h_1(i) h_2(i)}}\ , \tag{1}$$

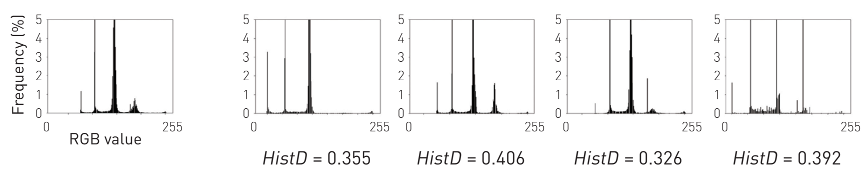
where $h_1(i)$ and $h_2(i)$ represent the frequency of the RGB value i in the histograms of GAN-generated layouts and human-designed layouts, respectively. The range of *HistD* is [0, 1], with 0 indicating identical histograms, and values less than 0.5 indicating a similar overall trend (Fig. 2).

3.1.2 Evaluation Metrics for Masterplan Rendering

The structural similarity index measure (*SSIM*) is a widely used tool for image similarity assessment, measuring the perceptual differences between two homogenous images (x, y) that have undergone different processes[23]. *SSIM* primarily evaluates the impact of luminance, contrast, and structural features on visual



	Human-designed	Mixed		Curvilinear		Polyline		Organic	
	BN	BN	Absolute BND	BN	Absolute BND	BN	Absolute BND	BN	Absolute BND
Green space	19	4	**15**	15	**4**	9	**10**	27	**8**
Path	1	1	**0**	5	**4**	1	**0**	3	**2**
Activity node	6	1	**5**	8	**2**	5	**1**	1	**5**
Small architectural feature	7	9	**2**	20	**13**	8	**1**	3	**4**
Waterscape	1	0	**1**	4	**3**	0	**1**	3	**2**



2. Diagram of block number distance and histogram distance methods.

perception based on two-dimensional grayscale images. In this study, *SSIM* is calculated to assess the differences between GAN-rendered and manually-rendered masterplans by landscape architects:

$$SSIM(x,y) = [l(x,y)]^{\alpha} \cdot [c(x,y)]^{\beta} \cdot [s(x,y)]^{\gamma}, \quad (2)$$

where $l(x,y)$ represents luminance, $c(x,y)$ represents contrast, and $s(x,y)$ represents structure. The parameters α, β, γ are greater than 0 and typically take the value of 1. The range of *SSIM* is [0, 1], with 1 indicating identical structures and 0 indicating completely different structures.

Additionally, the previously mentioned *HistD* can also be applied to evaluate the color distribution differences between two homogenous rendered masterplans, which was thus included for masterplan rendering evaluation.

3.2 Metrics for User Surveys

To evaluate whether GAN-generated layouts can visually mimic the human-designed ones and to understand professionals' color and texture preferences for renderings created by mainstream GAN models including BicycleGAN, it is necessary to include the practitioners in the surveys[24]. From September 1 to October 31, 2023, the research team conducted two online surveys using Sojump, targeting teachers, students, and professional designers in landscape architecture and related fields. The surveys were distributed to the School of Architecture and Planning at Hunan University, the School of Architecture at Tsinghua University, and the Beijing General Municipal Engineering Design & Research Institute Co., Ltd. Respondents were required to indicate their years of study or professional experience to ensure the representativeness and reliability of the results.

3.2.1 Questionnaire One

Questionnaire One aimed to conduct a Turing test for GAN-generated layouts and evaluate practitioners' acceptance of them. The questionnaire included 30 layouts, of which 16 were randomly selected from the Pix2Pix-generated layouts of the validation set, and 14 were layout redrawings of design from renowned firms or designers. Respondents were asked to identify the images they believe were generated by AI, with no limit on the number of selections (Fig. 3). These layouts are displayed in full screen on mobile devices with a resolution of 256 × 256 pixels, which is sufficient for estimation.

3.2.2 Questionnaire Two

The objective of Questionnaire Two was to evaluate the acceptance of renderings generated by mainstream GAN models. The questionnaire provided 30 rendered masterplans (10 groups, each containing 3 images from Pix2pix, CycleGAN, and BicycleGAN). Respondents were asked to determine whether the renderings met the standards for concept communication and to choose the best rendering from each group according to color and texture (Fig. 4). To facilitate detailed comparison, the masterplans were enlarged to 1,024×1,024 pixels and displayed in full screen on mobile phones.

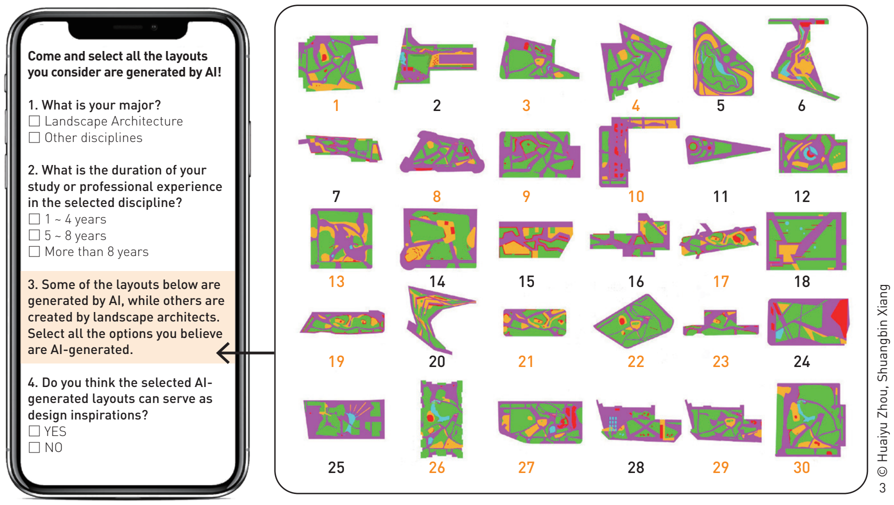
4 Evaluation Results

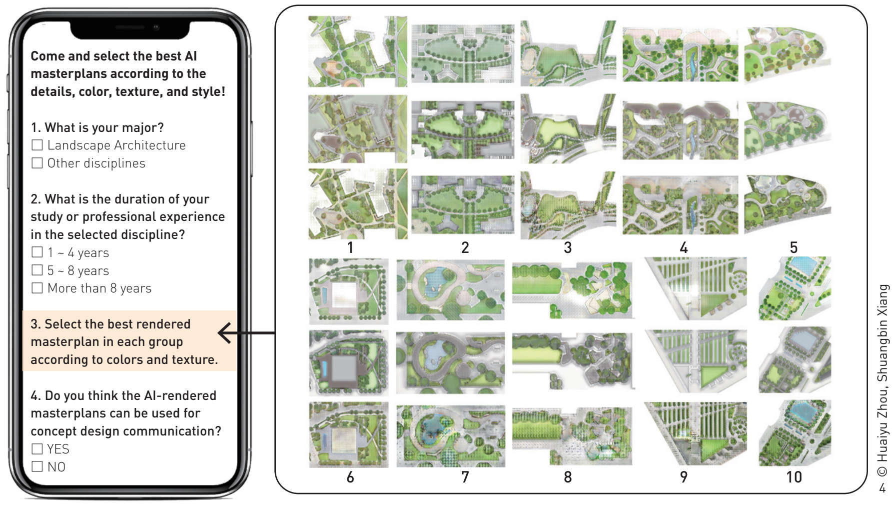
4.1 Image Analysis

4.1.1 Layout Generation Evaluation Results

A comparison between GAN-generated layouts and human-designed layouts reveals that both exhibit similar levels of diversity in land use BN statistically, with significant similarity in land area proportions.

1) According to the Quantile-Quantile plots and Shapiro–Wilk test results, the BN of five land use types in 340 GAN-generated layouts and human-designed layouts all follow a normal distribution. The average absolute BND calculation results (Table 1) show that for individual layouts, the differences in the number of the five land



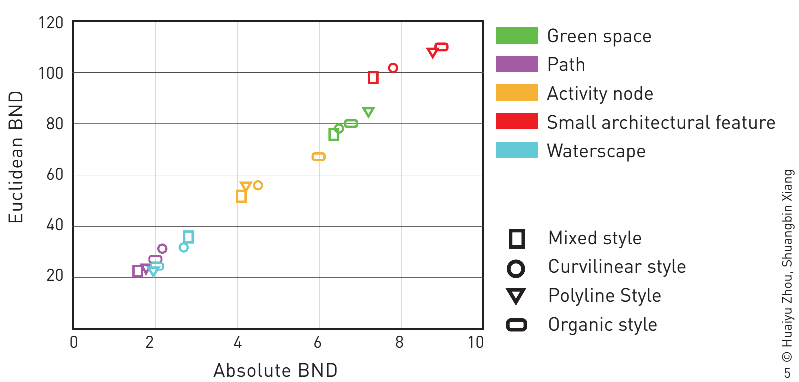


3. Online survey Questionnaire One (orange numbers represent GAN-generated layouts).

4. Online survey Questionnaire Two.

Table 1: Average absolute BND between GAN-Generated layouts and human-designed layouts

Land use type	Average BN of human-designed layouts	Average absolute BND of GAN-generated Layouts			
		Mixed	Curvilinear	Polyline	Organic
Green space	15.6	1.4	2.5	3.2	1.7
Activity node	6.5	2.1	2.5	3.2	2.1
Small architectural feature	12.0	4.3	4.3	4.8	3.0
Path	2.2	1.6	2.2	1.8	2.3
Waterscape	1.6	2.8	2.7	1.9	2.0



5. Cluster analysis of absolute BND and Euclidean BND between GAN-generated layouts and human-designed layouts.

use types between GAN-generated and human-designed layouts are all less than 5. The main differences lie in the number of small architectural features, suggesting that GANs and designers exhibit similar diversity in land use division.

2) To determine whether the differences in the number of layouts in the four styles' training sets would lead to significant differences in the BND results, this research further conducted a cluster analysis of the absolute BND and Euclidean BND for the four styles and five land-use types (Fig. 5). The results show that, for the same type of land-use block, the distributions of the four styles generally exhibit a clustering trend. This indicates that the four styles have strong similarities in land use division, and the different quantities in the training sets did not significantly affect the training results.

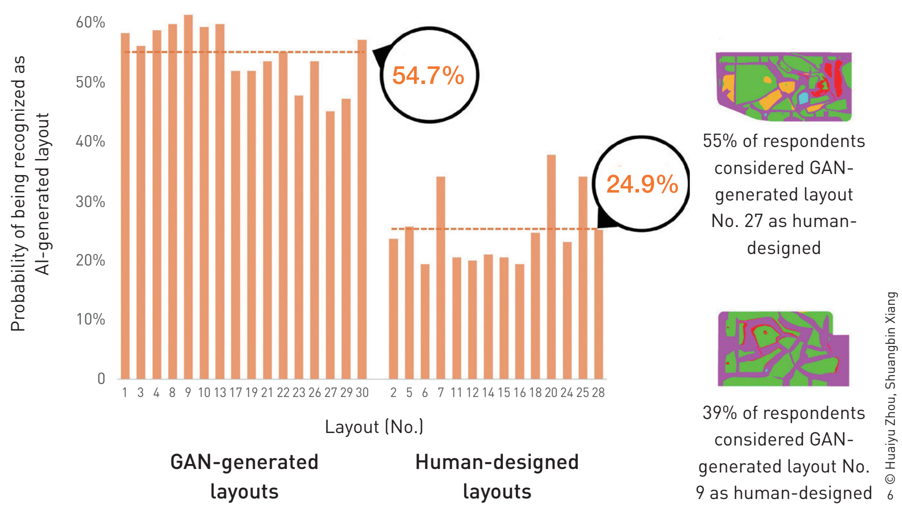
3) The average *HistD* values for the four styles are all less than 0.5: 0.41 (mixed style), 0.45 (curvilinear style), 0.41 (polyline style), and 0.43 (organic style), indicating a similar trend of the overall area proportion of different land use types in GAN-generated layouts to the human-designed layouts.

4.1.2 Masterplan Rendering Evaluation Results

The average *SSIM* and *HistD* values of the 50 rendered images were calculated, and the results are shown in Table 2: the average *SSIM* values for the warm-toned and cold-toned renderings by BicycleGAN are 0.786 and 0.790, respectively, which are close to 1; the average *HistD* values for the two tones are 0.391 and 0.406, respectively, which are less than 0.5. Additionally, when comparing the rendering results of Pix2Pix and CycleGAN, it was found that their average *SSIM* and *HistD* values are slightly different from

Table 2: Average ***SSIM*** **and** ***HistD*** **of GAN-rendered masterplans and human-designed masterplans**

GAN model		Average *SSIM*	Average *HistD*
BicycleGAN	Warm-toned	0.786	0.391
	Cold-toned	0.790	0.406
Pix2Pix	—	0.783	0.359
CycleGAN	—	0.795	0.436



6. Average probability of GAN-generated layouts and human-designed layouts being identified as AI generation output.

those of BicycleGAN, indicating a need to further investigate user preferences through surveys. Overall, the analysis results suggest that GAN-rendered masterplans are highly similar to those rendered by professional designers in terms of pixel distribution, structure, contrast, and luminance.

4.2 User Survey

4.2.1 Results of Questionnaire One

There are 192 valid responses received for Questionnaire One, of which 105 respondents had a background in landscape architecture, and the remainder came from related fields covering architecture, planning, and graphic design. Notably, 55% of respondents had over five years of professional experience, ensuring the reliability of the results. The findings indicate an average probability of 54.7% that the 16 GAN-generated layouts were identified to be AI generation outputs (Fig. 6), which was slightly higher than the probability of random guessing. Additionally, GAN-generated layouts had about a 45% probability of being mistakenly identified as designer-created layouts. Layout No. 27 (GAN-generated), deceived more than 55% of respondents. Furthermore, human-designed layouts had a probability of approximately 25% to be considered as GAN-generated. Overall, GAN-generated layouts can confuse some respondents, and about 70% of respondents believe that GAN technologies have the potential to assist concept design.

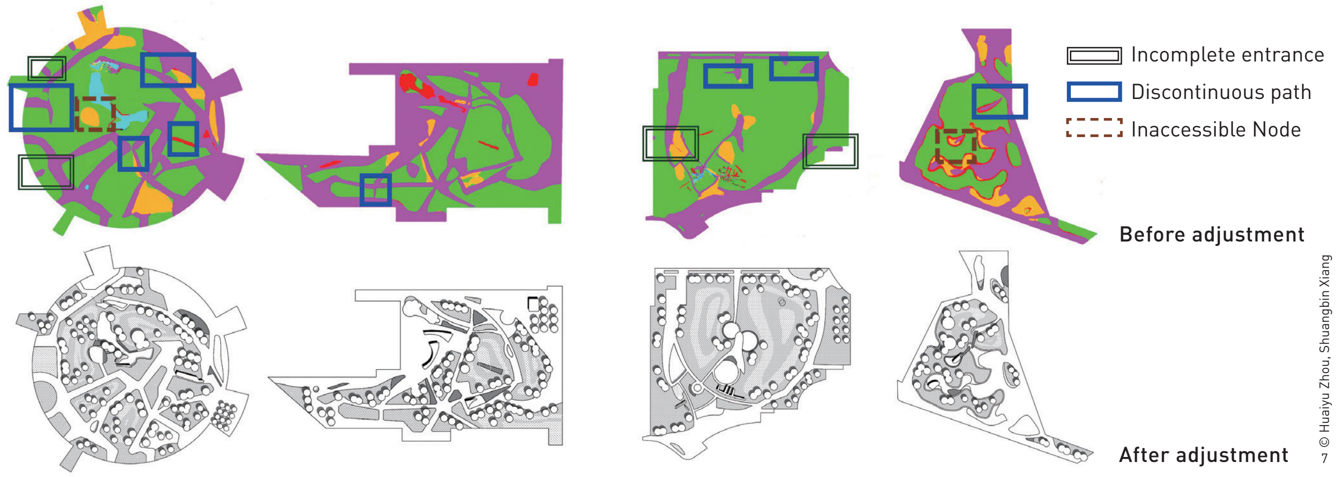
The study further communicated with respondents through phone calls, WeChat, emails, etc., to understand how they distinguish whether a layout is AI-generated or human-designed. It was found that unreasonable details in functional design severely compromise the visual realism of GAN-generated layouts. The study categorized the defects in GAN-generated layouts into three aspects (Fig. 7): 1) incomplete entrance, which is too small or lacks connections to internal roads, preventing access to the site; 2) discontinuous path, which may be interrupted by green spaces or outdoor facilities and impedes movements; 3) inaccessible node, which is the isolated spaces that cannot be reached via paths. The study further quantified the average occurrence of these defects in the 340 GAN-generated layouts. As shown in Table 3, the issue of discontinuous path is the most prominent.

4.2.2 Results of Questionnaire Two

Questionnaire 2 received a total of 422 valid responses, among which 233 respondents (55%) had a background in landscape architecture, and 155 respondents (37%) had more than five years of professional experience. The results indicate that 91% of respondents believe that the quality of GAN-rendered masterplans meets the requirements for conceptual design refinement and communication. Regarding the renderings generated by BicycleGAN, Pix2Pix, and CycleGAN, 47% of respondents considered BicycleGAN to perform the best in color and texture, suggesting that the model upgrade from CycleGAN to BicycleGAN was well-received by users.

5 Conclusions and Discussion

This study introduced image analysis and user survey metrics to evaluate the applicability of GAN generation methods, aiming to fill the gap that existing research primarily focuses on model training methods but lacks post-evaluation. This research provides an accessible evaluation framework for “image-to-image” generative design research. The image analysis results show that both the



7. Examples of three kinds of defects in GAN-generated layouts and adjustment solutions from landscape architects.

diversity similarity of land use distribution in GAN-generated layouts and human-designed layouts and the similarity between GAN-rendered masterplans and those rendered by designers achieved a high level. The user survey results indicate that GAN-generated layouts are difficult to distinguish from human-designed ones, with their rendered colors and textures accepted by landscape architects. Referring to the explanation of the paradigms of landscape design thinking by Yufan Zhu[25], the inherent rationality of the "image-to-image" model can be demonstrated from a historical perspective. Experienced the limited styles of classical gardens, the late 19th and early 20th-century artists such as Piet Cornelis Mondrian, Jean Arp, and Roberto Burle Marx, vividly illustrated the empathy between painting and landscape design—abstract forms guide the spatial structure of landscape[25]. As an interdisciplinary field of engineering and art, modern landscape architecture largely relies on intuition, experience, and emotion, making its forms characterized by nonlinear processes and become challenging to quantify. Artificial neural networks share similarities with human neurons, and efforts are being made to train them to learn how pioneer designers applied limitless styles to diverse sites. Although the GAN generation model involves black-box processes, this study provides quantitative support for the rationality of its internal logic.

This research has several limitations. First, it did not include an ethical evaluation of GAN generation methods. Currently, GAN generation methods have raised discussions in professional ethics about how to maintain designers' creativity and in design education about how to reasonably integrate AI technology in study. Generally, it is necessary to address functional requirements in design according to regional and environmental contexts. However, GAN generation methods often lack an understanding of form symbols under complex historical and cultural influences. The questionnaires lacked focus on the originality of GAN generation methods, and future research needs to collect user opinions on ethical issues. Second, the established evaluation framework does not consider the diversity of GAN-generated layouts and data bias. The AI outputs are significantly influenced by their training data. Currently, the diversity of landscape masterplan datasets is severely lacking, especially with few for classical Chinese gardens compared to modern landscape projects. Application of these tools can lead to homogenized design results. Future research needs to explore how to maintain the diversity of design. Taking design courses as an example, students who have not yet established a comprehensive knowledge system may lack the ability to judge the quality of datasets when using AI tools. Simply using GAN-generated layouts to complete coursework could limit the process of knowledge acquisition and design skill development. Furthermore, while the Pix2Pix–BicycleGAN workflow evaluated in this study is representative, it does not reflect the latest technological iterations. Future research could explore customized GAN models for specific regions or types of landscape design (e.g., the classical Chinese gardens, the Western modern landscapes), incorporating

Table 3: Summary of defects in GAN-generated layouts

Defect	Average number			
	Mixed	Curvilinear	Polyline	Organic
Incomplete entrance	2.1	2.5	2.2	4.0
Discontinuous path	3.3	4.8	5.8	5.0
Inaccessible node	1.6	2.2	1.8	2.0

more data with regional features for model training and developing algorithms that can identify and emphasize these features[26].

In addition, evidence-based design challenges GAN generation methods due to its relatively low interpretability[27][28]. Apart from morphology, the scientific thinking of "design with nature" requires the integration of various factors (e.g., topography, soil, runoff, and vegetation) to justify design decisions. The application of physical models and monitoring technologies, including hydrodynamic modeling and Internet of Things (IoT) technology[29]~[31], helps increase the interpretability of the design. Therefore, connecting the morphological expressions generated by GAN models with the quantitative analyses (e.g., physical models) should be overcome for deeply integrating AI into design disciplines. The research team from South China Agricultural University has already attempted to couple GAN generation methods with evidence-based health design strategies to design age-friendly gardens[32][33]. As the diversity of GAN-generated layouts increases, future utilization of multi-optimization algorithms to screen and improve the layouts will help enhance the scientific decision of design[34]. With the ongoing updates of generative algorithms, there are opportunities to gradually integrate physical models and optimization algorithms with AI models, significantly improving the interpretability and applicability of GAN generation methods.

ACKNOWLEDGEMENTS

The authors extend their gratitude to the faculty and students of the School of Architecture and Planning at Hunan University, the School of Architecture at Tsinghua University, as well as the designers from the Beijing General Municipal Engineering Design & Research Institute Co., Ltd. for participating in the survey. The authors also thank assistant professor Ran Yi from Shanghai Jiao Tong University and tenured associate professor Hailong Liu from Tsinghua University for their assistance and support to this research.

ELECTRONIC SUPPLEMENTARY MATERIAL

Supplementary material is available in the online version of this article at https://doi.org/10.15302/J-LAF-1-020094.

Competing interests | The authors declare that they have no competing interests.

REFERENCES

[1] Bao, R. (2019). Research on intellectual analysis and application of landscape architecture based on machine learning. *Landscape Architecture*, 26(5), 29–34.

[2] Zhao, J., & Cao, Y. (2020). Review of artificial intelligence methods in landscape architecture. *Chinese Landscape Architecture*, 36(5), 82–87.

[3] Zhao, J., Chen, R., Hao, H., & Shao, Z. (2021). Application progress and prospect of machine learning technology in landscape architecture. *Journal of Beijing Forestry University*, 43(11), 137–156.

[4] Huang, W., & Zheng, H. (2018). Architectural Drawings Recognition and Generation Through Machine Learning. In: P. Anzalone, M. D. Signore, & A. J. Wit (Eds.), *Proceedings of the 38th Annual Conference of the Association for Computer Aided Design in Architecture* (pp. 18–20). ARCADIA.

[5] 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. In: A. Vedaldi, H. Bischof, T. Brox, & J.-M. Frahm (Eds.), *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I* (pp. 162–177). Springer.

[6] Newton, D. (2019). Deep generative learning for the generation and analysis of architectural plans with small datasets. *Proceedings of 37th eCAADe and 23rd SIGraDi Conference*, (2), 21–28.

[7] Chen, M., Zheng, H., & Wu, J. (2022). Computational design of multi-functional system based on generative adversarial networks: Taking the layout generation of Vocational and Technical College as an example. *Architectural Journal*, (S1), 103–108.

[8] Lin, W. (2020). *Research on automatic generation of primary school schoolyard layout based on deep learning* [Master's thesis]. South China University of Technology.

[9] Sun, C., Cong, X., & Han, Y. (2021). Generative design method of forced layout in residential area based on CGAN. *Journal of Harbin Institute of Technology*, 53(2), 111–121.

[10] Zhang, T. (2020). *Experiments on generation of the arrangement of residential groups based on deep learning* [Master's thesis]. Nanjing University.

[11] Zhou, H., & Liu, H. (2021). Artificial intelligence aided design: Landscape plan recognition and rendering based on deep learning. *Chinese Landscape Architecture*, 37(1), 56–61.

[12] Qu, G., & Xue, B. (2022). Generative design method of landscape functional layout in residential areas based on Conditional Generative Adversarial Nets. *Low Temperature Architecture Technology*, 44(12), 5–9, 14.

[13] Chen, R., & Zhao, J. (2023). Generation and design feature recognition of landscape architecture scheme based on style-based generative adversarial network. *Landscape Architecture*, 30(7), 12–21.

[14] Zhao, G. (2023). *Research on application of generative model in landscape design* [Master's thesis]. Shanxi University.

[15] Zhou, W. (2023). *Design research on pocket park plan layout generation based on deep learning* [Master's thesis]. Chongqing Jiaotong University.

[16] Huang, Y., & Zhou, Y. (2023). Exploration on the generative architecture design method with AIGC technology: A case of the overall design process of generating architectural image with prompt as a key word. *Urbanism and Architecture*, 20(15), 202–206, 213.
[17] Chen, J., Shao, Z., & Hu, B. (2023). Generating interior design from text: A new diffusion model-based method for efficient creative design. *Buildings*, 13(7), 1861.
[18] Turrin, M., von Buelow, P., & Stouffs, R. (2011). Design explorations of performance driven geometry in architectural design using parametric modeling and genetic algorithms. *Advanced Engineering Informatics*, 25(4), 656–675.
[19] Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image Translation With Conditional Adversarial Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1125–1134). IEEE.
[20] Zhu, J.-Y., Zhang, R., Pathak, D., Darrell, T., Efros, A. A., Wang, O., & Shechtman, E. (2017). Toward Multimodal Image-to-image Translation. *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 465–476). Curran Associates Inc.
[21] Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-image Translation Using Cycle-Consistent Adversarial Networks. *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2223–2232). IEEE.
[22] Cha, S.-H., & Srihari, S. N. (2002). On measuring the distance between histograms. *Pattern Recognition*, 35(6), 1355–1370.
[23] Hore, A., & Ziou, D. (2010). Image Quality Metrics: PSNR vs. SSIM. *2010 20th International Conference on Pattern Recognition* (pp. 2366–2369). IEEE.
[24] Geman, D., Geman, S., Hallonquist, N., & Younes, L. (2015). Visual turing test for computer vision systems. *Proceedings of the National Academy of Sciences*, 112(12), 3618–3623.
[25] Zhu, Y. (2022). Disordering and redirecting: Paradigm of design thinking in contemporary landscape architecture. *World Architecture*, (11), 36–37.
[26] Jiang, F., Ma, J., Webster, C. J., Li, X., & Gan, V. J. (2023). Building layout generation using site-embedded GAN model. *Automation in Construction*, (151), 104888.
[27] Li, P., Liu, B., & Gao, Y. (2018). An evidence-based methodology for landscape design. *Landscape Architecture Frontiers*, 6(5), 92–101.
[28] Yang, Y., & Lin, G. (2020). The development, connotations, and interests of research on landscape performance evaluation for evidence-based design. *Landscape Architecture Frontiers*, 8(2), 74–83.
[29] Zhou, H., Jiang, H., & Liu, H. (2021). Process visualization and performance evaluation of stormwater management in landscape projects based on IoT online monitoring. *Chinese Landscape Architecture*, 35(10), 29–34.
[30] Zhou, H., & Liu, H. (2021). IoT-based operational information management for built landscape projects: From vacancy to approaches. *Landscape Architecture Frontiers*, 9(2), 83–95.
[31] Zhou, H., Li, R., Liu, H., & Ni, G. (2023). Real-time control enhanced blue-green infrastructure towards torrential events: A smart predictive solution. *Urban Climate*, (49), 101439.
[32] Li, H., Zhang, Z., Liu, K., Chen, W., Wei, W., Liu, X., Xie, J., Zhang, M., Huang, Z., Zhong, M., Cai, C., Huang, X., Hou, Y., Lin, X., Yu, S., Fang, Y., & Feng, X. (2023, November 25). *Toward dynamic optimization: Combining AI and EBHDL for the elderly*. American Society of Landscape Architects.
[33] Chen, C., Li, H., Hou, Y., & Liu, J. (2023). Application progress of computer vision in the research on relationship between landscape and health. *Landscape Architecture*, 30(01), 30–37.
[34] Liu, H., Jin, C., & Yang, Y. (2023). Study on the programming language and its organicity of architectural generative design. *Urbanism and Architecture*, 20(5), 182–186.

https://doi.org/10.15302/J-LAF-1-020094 收稿日期 | 2023-11-26 录用日期 | 2024-02-06 中图分类号 | TU986.2, TP181 文献标识码 | A

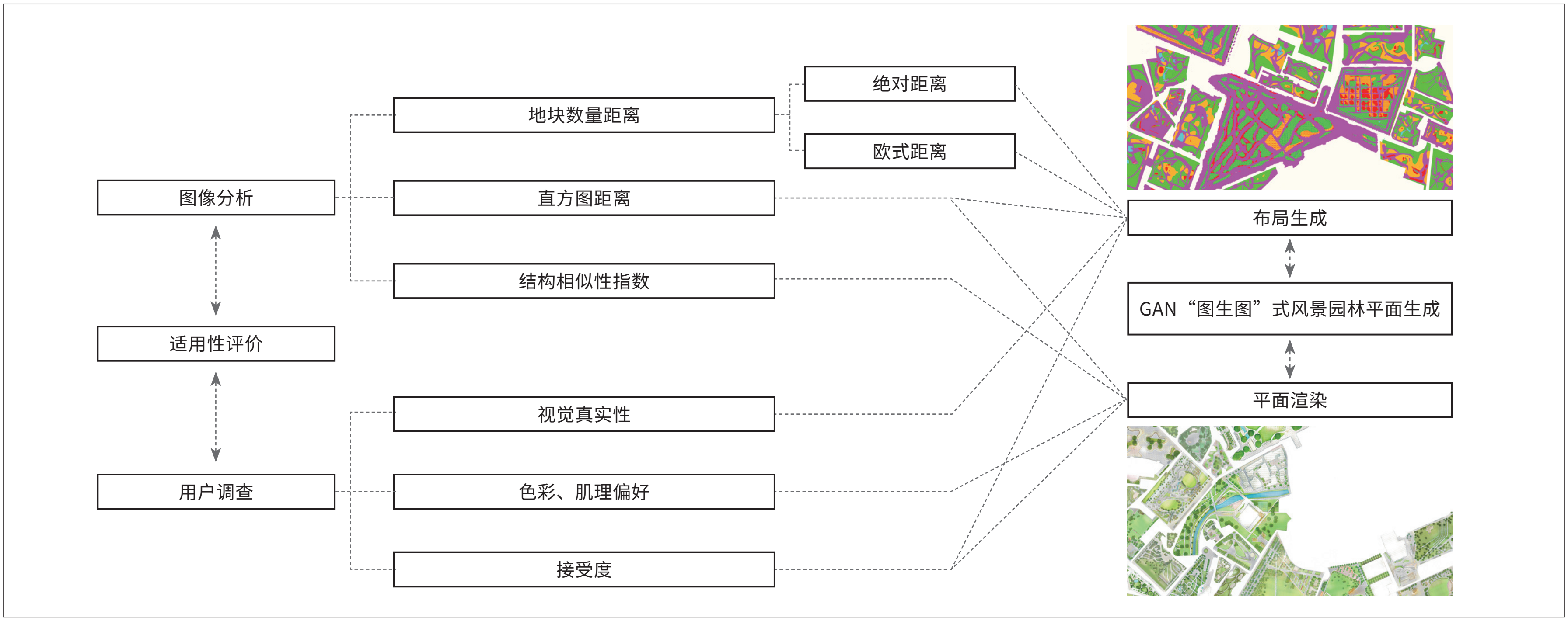
人工智能“图生图”式景观平面生成技术的适用性评价与反思

周怀宇[1,2,*]，向双斌[2,3]

1 湖南大学建筑与规划学院建筑系，长沙 410082
2 清华大学建筑学院景观学系，北京 100084
3 北京市市政工程设计研究总院有限公司建筑院景观室，北京 100088

*通讯作者
地址：湖南省长沙市岳麓区湖南大学建筑与规划学院
邮编：410082
邮箱：zhouhy@hnu.edu.cn

图文摘要



摘要

人工智能（AI）图像生成技术正在改变景观设计中的传统工作模式，其中，“图生图”式生成对抗网络（generative adversarial network，GAN）技术具备辅助方案设计的潜能，因此面向用户端对其展开技术适用性评价研究对于优化工具选择、提升设计效率尤为重要。本研究旨在借助图像分析和用户调查方法，评估GAN生成方法生成结果的质量、与设计工作对接的有效性，以及景观设计师对图像生成结果的接受度。研究以Pix2Pix－BicycleGAN工作流中布局生成与平面渲染两项任务为评价对象，建立了基于地块数量的绝对/欧式距离、直方图距离、结构相似性指数等图像分析指标；针对GAN生成结果的视觉真实性和色彩肌理偏好开展了两项在线用户问卷调查。结果显示，GAN生成布局与真实布局相似性高，GAN渲染平面能够满足概念方案呈现要求、用户接受度好。最后，本文探讨了GAN生成方法的内在合理性及其在行业伦理及数据偏见方面的局限性，反思现阶段连接AI辅助设计与循证设计之间的技术空缺。

关键词

景观设计学；图像生成；生成对抗网络；人工智能辅助设计；适用性评价；景观平面

文章亮点

- 定量评价GAN“图生图”式景观设计平面生成技术的适用性
- 图像分析显示GAN生成布局与真实布局高度相似
- 用户调查显示GAN渲染平面视觉真实性较高且从业人员对其接受度高
- 指出现阶段GAN生成方法的内在合理性及其与循证设计之间的技术空白

编辑 高雨婷，王颖
翻译 高雨婷

1 引言

在近年来兴起的生成式人工智能（AI）热潮中，迅速发展、升级的图像生成技术和制图工具不断冲击着传统景观设计行业的工作模式[1]~[3]。因而景观设计师迫切需要厘清图像生成与景观设计之间的关系，探索新的行业增长点与研究机遇。作为图像生成技术的用户端，景观设计师对该技术的适用性评价能够协助分析其潜在影响、优化工具选择，进而提升设计效率。目前可以对接景观设计工作流的图像生成技术主要被应用于平面生成及效果图渲染两方面。

平面生成的相关研究主要基于“图生图”式生成对抗网络（generative adversarial network，GAN）开展。这类工具以建筑户型平面生成[4]~[6]为起点，目前已经发展至建筑排列方式与体块关系的生成[7]~[10]。近年来，景观设计领域也开启了平面生成的研究：周怀宇等建立了平面标注数据集，基于CycleGAN①实现了景观平面的识别和渲染[11]；曲广滨等基于CGAN②生成了满足设计规范要求的居住区景观功能平面[12]；陈然等利用StyleGAN2③生成了多样化的设计方案，并发现GAN模型能够识别和提取植被、水域、铺装和路网等的高维抽象特征[13]；赵冠婕耦合利用Pix2Pix④和Stable Diffusion模型，探讨了小尺度景观间的自动化设计流程[14]；周韦世则利用Pix2Pix来生成小尺度的城市口袋公园平面布局[15]。尽管现有研究已较为深入地探讨了训练原理、数据集和生成方法，但仍存在以下问题——缺乏公开可获取的景观平面数据集，训练数据丰富性较低；可生成平面的尺度有限，主要适用于中小型绿地；针对GAN所生成平面的系统化定量评价较少，缺乏便于操作的评价指标；针对用户端开展的调查较少，难以获取使用评价。

效果图生成的相关研究与应用主要围绕Midjourney和Stable Diffusion两大“文字生图”（text to image）工具开展。Midjourney模型可通过各类提示词在网页端生成精细度与真实度较高的人视图或鸟瞰图，操作较为便捷。相比之下，开源的Stable Diffusion模型除了可以通过关键词生成图片外，还具备“图生图”（image to image）和“模型生图”（model to image）的自由训练功能，方便设计师添加限制条件，因而更受欢迎。目前，基于Stable Diffusion的建筑形体构思和建模工作流已经初步形成[16][17]。

随着技术的迭代更新，设计工作流经历了从手绘到CAD制图、再到借助Grasshopper的参数化设计的发展，设计师通过在实践中融合新技术、积极开展评价，推动了新工作流的出现[18]。本研究关注基于GAN的景观平面生成方法（后文简称“GAN生成方法”），从景观设计师的视角综合评估其技术适用性，以期为设计师在选择工具时提供决策依据。在“图生图”工作中，GAN生成方法的核心在于提升制图效率，能够减少设计师在多方案比选和反复渲染过程中花费的时间[11]。因此，本研究旨在借助图像分析和用户调查方法，评估GAN生成方法生成结果的质量、与设计工作对接的有效性，以及景观设计师对图像生成结果的接受度。

2 评价对象：Pix2Pix-BicycleGAN工作流

本研究着眼于Pix2Pix－BicycleGAN景观平面生成工作流中两项关键任务——布局生成与平面渲染——的适应性评价。GAN生成的布局类似设计教学中的功能泡泡和平面草图，代表一种直观、简洁的思维方式，是设计迭代和调整的基础；GAN渲染图则为布局中抽象的形态添加了色彩和肌理细节而使其更具可读性。任务实现工具Pix2Pix[19]是GAN领域应

① CycleGAN，即循环生成对抗网络，能够在两个不同的图片示例之间进行无监督的图像转换，而无需成对地训练数据，适用于风格迁移。
② CGAN，即条件生成对抗网络，它在图像生成过程中引入了额外的条件或标签，使得生成结果不仅受到随机噪声的影响，还受到特定条件限制。
③ StyleGAN，即基于风格的生成对抗网络，能够通过对风格向量的训练增强所生成图像的多样性，在图像细节和大规模数据集处理方面表现出色。
④ Pix2Pix，即“图生图”生成对抗网络，能够根据输入图像生成对应的输出图像，如将黑白转换为彩色、将草图转换为照片等，在图像修复合成任务上表现出色。

用较为广泛的模型，而BicycleGAN[20]是CycleGAN[21]的改进模型，通过引入额外的变量并增加约束条件，提高了模型处理多模态数据和高分辨率图像的性能，同时支持多种渲染结果输出。由于数据集中获取与标注的平面类型有限，这一工作流目前主要适用于中小尺度的景观场地[11][13]-[15]（图1）。

2.1 生成多种样式风格的布局

通过向Pix2Pix模型输入场地范围，可以生成多种风格（混合、曲线、折线）且包含不同用地类型的场地布局，包含绿地、路径、活动节点、构筑物小品和水景（图1）。对GAN布局生成的评价围绕生成的用地地块布局与真实布局（由人类设计师完成的设计方案布局）的相似性和视觉真实性开展。

在本研究中共收集2 725张真实景观平面图，其中混合、曲线、折线、有机混合训练集分别为2 670、916、770、954张⑤，预留用于评估生成效果的验证集85张。依据景观平面生成PNG或JPEG格式的场地布局，以黑色填充图块表示场地范围，以蓝色圆圈标注出入口位置（其大小表示出入口等级）。基于4种样式风格，共得到340个（85×4）GAN生成布局用于后续评价。设计师在比较多个GAN生成布局后，继续调整设计形态、补充种植设计、细化用地地块划分，依据项目需求并结合个人经验形成更为精准的地块布局，并将其作为平面渲染任务的输入。

2.2 渲染多种色彩肌理的平面

设计师将调整后的布局输入BicycleGAN中，可获取不同色彩肌理的渲染平面[20]，方便与业主快速沟通设计思路。该任务的评价主要围绕GAN渲染平面与人工渲染平面（如手绘填色或利用Adobe Photoshop等工具填色）的相似性及用户色彩肌理偏好开展。数据集共包含景观平面325张，其中训练集300张，验证集25张[11]。由于Bicycle会生成多种渲染结果，每张布局挑选暖色调、冷色结果各一张，评价总量为50张。笔者过往研究[11]中使用相同布局输入生成的Pix2Pix和CycleGAN的平面渲染结果也一并纳入本研究的图像分析与用户调查中⑥。

3 评价方法

对GAN生成布局与真实布局、GAN渲染平面与人工渲染的相似性的评价聚焦于图片本身的特征，适宜采用图像分析方式，而对于生成布局的视觉真实性和渲染偏好的评价则适宜采用用户调查的方式，因此本研究耦合图像分析与用户调查建立评价指标体系。

⑤ 混合训练集为曲线、折线、有机三个训练集的集合。

⑥ 具体模型训练方法详见参考文献[11]。

3.1 图像分析指标

3.1.1 生成布局评价指标

（1）地块数量距离

所生成五类用地的地块数量（BN）能够最直观地反映GAN生成布局的形态多样性，相应的地块数量距离（BND）可用于评估340张由Pix2Pix生成的验证集布局和真实布局的差异。其中，BND评价包含绝对BND和欧氏BND两项指标的计算。本研究通过绝对距离比较单一样式风格下生成布局中各类用地BN（人工计数同一RGB值贯通的完整色块）与真实布局之间的差值。同时，为补充分析各样式风格的训练集布局数量的不同造成的影响，本研究通过绝对BND与欧式BND的聚合分析比较了四类样式风格之间地块划分聚集程度的差异，并以聚合图呈现两组数据的中点聚集区。

（2）直方图距离

图像直方图可显示图像中不同RGB像素的频率分布，直方图距离（*HistD*）则是衡量两幅图像之间像素分布差异的关键指标[22]。基于GAN生成布局中的RGB值和用地类型一一对应的关系，*HistD*能够有效评估GAN生成布局与真实布局在用地地块划分与面积比例上的差异。本研究通过巴氏距离来量化两张归一化直方图之间的距离：

$$HistD = \sqrt{1-\sum_i \sqrt{h_1(i)h_2(i)}}\ ,\qquad (1)$$

式中，$h_1(i)$和$h_2(i)$代表计算GAN生成布局与真实布局两类直方图中第i个RGB值的出现频率；*HistD*的取值范围为[0,1]，取值为0表示两个直方图相同，取值小于0.5代表二者总体呈现相似趋势（图2）。

3.1.2 平面渲染评价指标

结构相似性指数（*SSIM*）是一种广泛使用的图像相似性度量工具，可以评估两幅经过不同处理加工的同源图像（x, y）之间的感知差异[23]。*SSIM*主要基于二维灰度图像来评估亮度、对比度和结构特征对视觉的影响。本研究通过计算*SSIM*来评估渲染平面与景观设计师人工渲染平面的差别：

$$SSIM(x,y) = [l(x,y)]^{\alpha}\cdot[c(x,y)]^{\beta}\cdot[s(x,y)]^{\gamma}\ ,\qquad (2)$$

式中，$l(x, y)$代表亮度、$c(x, y)$代表对比度，$s(x, y)$代表结构，α、β、γ为大于0的参数，常取值1。*SSIM*的取值范围为[0,1]，其中，1表示两幅图像具有相同的结构，0则表示完全不同。

此外，上述*HistD*指标亦可用于评估两幅同源渲染平面的色彩分布差异。因此，*HistD*也被并纳入平面渲染评价指标。

3.2 用户调查指标

评价GAN生成布局在视觉上能否以假乱真，同时了解从业人员对BicycleGAN及其他多种主流GAN模型渲染图的色彩肌理偏好，需要专业人士共同评估[24]。研究团队于2023年9月1日至10月31日，面向景观设计及相关领域的教师、学生和职业设计师发布了两项问卷星在线调查问卷。问卷主要被投放到湖南大学建筑与规划学院、清华大学建筑学院及北京市市政工程设计研究总院，同时要求受访者选择其求学或从业年限以确保结果的代表性与可靠性。

3.2.1 问卷1

问卷1旨在对GAN生成布局进行图灵测试，并评估从业人员对GAN生成布局的接受度。问卷中共涉及30张平面布局，其中16张随机抽取自验证集的Pix2Pix自动生成布局，14张由知名事务所或大师创作方案的布局改绘，受访者需要从中选出他们所认为的由AI生成的图片，问卷并未设置最多可选数量限制（图3）。这些布局以256×256的像素在手机上全屏呈现，可以满足判别要求。

3.2.2 问卷2

问卷2旨在判断从业人员对几类主流GAN模型渲染图接受度的差异。问卷提供了30张渲染平面图（10组、每组3张，分别来自Pix2pix、CycleGAN和BicycleGAN），要求受访者判断渲染图是否达到在概念设计阶段用于方案交流的标准，并依据色彩和肌理选择每组中效果最佳的平面（图4）。为方便用户细致比较图片渲染细节，平面图被局部放大至1 024×1 024像素在手机上全屏呈现。

4 评价结果

4.1 图像分析

4.1.1 生成布局评价结果

总体对比GAN生成布局与真实布局发现，两者在图形统计意义上的BN多样性水平接近，地块面积比例相似性突出。

1）经Quantile-Quantile图和Shapiro-Wilk检验，340张GAN生成布局与真实布局的五类用地BN均符合正态分布。由绝对BND平均值计算结果（表1）可知，单张布局中，GAN生成的五类用地BN与真实布局的差别均小于5个，主要差别体现在小品构筑物的数量上，这表明GAN与设计师在用地划分时表现出的多样性较为相似。

2）为了确定4类样式训练集的布局数量差异是否会导致BND结果的显著不同，进一步对四类风格、五类用地地块的绝对BND与欧氏BND进行聚合分析（图5）。结果显示，就相同用地地块类型而言，四类样式的分布总体呈现聚拢趋势，表明4种样式风格在用地划分上也具有较强的相似性，进而可知训练集数量的不同并没有显著影响训练结果。

3）混合、曲线、直线、有机四类样式风格的平均*HistD*值分别为0.41、0.45、0.41、0.43，均小于0.5，意味着GAN生成布局对不同用地类型划分的面积比例总体与真实布局呈现接近趋势。

4.1.2 平面渲染评价结果

计算50张渲染图的平均*SSIM*和*HistD*值，结果如表2所示：BicycleGAN暖色、冷色调的平均*SSIM*分别为0.786和0.790，接近1；两种色调的平均*HistD*值分别为0.391和0.406，小于0.5。同时，比较Pix2Pix和CycleGAN的渲染结果，发现两者的平均*SSIM*和*HistD*值与BicycleGAN相差较小，需要结合调查进一步探讨用户偏好。总体来说，分析结果表明GAN渲染平面在像素分布、结构、对比度和亮度方面与职业设计师绘制的渲染图高度相似。

表 1：GAN 生成布局与真实布局的绝对 BND 平均值统计

用地地块类型	真实布局BN 平均值（个）	生成布局绝对 BND 平均值（个）			
		混合	曲线	直线	有机
绿地	15.6	1.4	2.5	3.2	1.7
活动节点	6.5	2.1	2.5	3.2	2.1
小品构筑物	12.0	4.3	4.3	4.8	3.0
路径	2.2	1.6	2.2	1.8	2.3
水景	1.6	2.8	2.7	1.9	2.0

表 2：GAN 渲染平面与设计师渲染平面的平均 *SSIM* 与 *HistD*

GAN 模型		平均 *SSIM*	平均 *HistD*
BicycleGAN	暖调	0.786	0.391
	冷调	0.790	0.406
Pix2Pix	—	0.783	0.359
CycleGAN	—	0.795	0.436

表 3：GAN 生成布局中的缺陷总结

缺陷类型	平均数量（个）			
	混合	曲线	直线	有机
入口不完整	2.1	2.5	2.2	4.0
道路不连贯	3.3	4.8	5.8	5.0
节点不可达	1.6	2.2	1.8	2.0

4.2 用户调查

4.2.1 问卷1结果

问卷1共收到192份有效回复，其中105位受访者具有景观设计专业背景，其余均来自建筑、规划、平面设计相关领域，55%的受访者有5年以上的从业经历，保证了结果的可靠性。结果显示，16张GAN生成布局被识别为AI生成的平均概率为54.7%（图6），略高于随机猜测的概率。而GAN生成布局有约45%的几率被从业人员错认为是设计师创作的布局。编号27的GAN生成布局欺骗了55%以上的受访者。同时，设计师创作的真实布局有约25%的概率被判定为GAN生成。总体而言，GAN生成布局可以使一些受访者感到迷惑，同时，约70%的受访者认为GAN技术有助力方案设计的潜力。

研究进一步通过电话、微信、邮件等形式与受访者交流如何辨别布局是由AI生成还是设计师绘制，发现功能设计中不合理的细节会严重破坏GAN方案的视觉真实性。研究将GAN布局缺陷分为三类（图7）：1）入口不完整，即入口过小或缺乏与内部道路的连接，而导致使用者难以进入场地；2）道路不连贯，即内部道路被绿地或室外设施打断而形成通行障碍；3）节点不可达，即存在孤立的、无法通过路径到达的空间。研究进一步统计了340张GAN生成布局中3类缺陷出现的平均次数。由表3所示结果可知，道路不连贯的问题最为明显。

4.2.2 问卷2结果

问卷2共收到422份有效问卷，受访者中233位（55%）具有景观设计专业背景，155（37%）位具有5年以上的从业经历。结果显示，91%的受访者认为GAN平面渲染的质量可以满足概念设计阶段的方案推敲与沟通。针对BicycleGAN、Pix2Pix和CycleGAN渲染图，47%的受访者认为BicycleGAN在色彩和肌理效果上表现最佳，这也间接表明由CycleGAN到BicycleGAN的模型升级得到了用户的认可。

5 结论与讨论

本文通过引入图像分析及用户调查指标来评估GAN生成方法的技术适用性，旨在填补现有研究主要关注训练方法而缺少后期评估的空白，为“图生图”生成式设计研究提供易于操作的评价框架。图像分析结果显示，GAN生成布局与真实布局的用地分布多样性、渲染平面图与设计师渲染平面图的相似性均达到了较高的水平；用户调查结果显示，GAN生成布局具有较强的迷惑性、真假难辨，且渲染色彩和肌理得到了景观设计师的认可。参考朱育帆对景观设计思维范式的阐释[25]，可以从学科历史观来补充论证“图生图”模式的内在合理性。经历古典时代有限的样式园林，皮特・科内利斯・蒙德里安、让・阿尔普、罗伯特・布雷・马尔克斯等19世纪末、20世纪初的艺术家们生动诠释了绘画与景观设计之间的同理性——由抽象图形指导场地景观空间结构[25]。作为工程与艺术的交叉学科，现代景观设计最难以量化的就是形态，它往往依赖于直觉、经验、情感等个人意志，且其生成过程具有非线性特征。而人工神经网络与人类神经元具有相似性，人们正尝试训练它学习先驱们如何将无限的样式赋予多样化的场地。即使GAN生成方法模型内部存在较多的黑箱过程，但本研究为其内在逻辑的合理性提供了定量化支撑。

本研究的局限性主要体现在以下几个方面。首先，研究未涉及对GAN生成方法的伦理评价。目前，GAN生成方法已经在行业伦理和设计教育方面引发了关于如何保持设计师创造力及如何在设计教育中合理融入AI技术的讨论。通常而言，设计需要基于特定的地域环境背景来实现功能需求，而GAN生成方法往往对复杂历史、文化因素影响下的形式符号缺乏理解。本研究的问卷调查缺少对GAN生成方法原创性的关注，需在未来的研究中补充收集用户对伦理问题的看法。其次，在本研究建立的评价框架中，未纳入对GAN生成布局多样性和训练数据偏见问题的考量。AI工具输出的内容受其训练数据影响显著，而目前景观平面数据集多样性严重不足，特别是中国古典园林平面布局的数据明显少于现代景观作品，盲目应用会导致设计成果的同质化，未来亟需探索如何避免潜在的设计多样性缺失。以设计课教学为例，尚未建立完善知识体系的学生在使用AI工具时可能缺乏对数据集质量的判断能力，学生“套用”GAN生成布局来完成课程作业的模式将限制其知识的获取和设计能力的提升。再者，本研究评价的Pix2Pix-BicycleGAN工作流虽然具有一定的典型性，但尚不能代表最前沿的技术迭代。在未来的研究中，可探索针对特定区域或类型景观设计（如中国古典园林与西方现代景观）的定制化GAN模型，在模型训练过程中融入具有更多地域特征的数据，以及开发能够辨识并强调这些特征的算法[26]。

此外，GAN生成方法较低的可解释性使其面临着来自循证设计的挑战[27][28]。形态只是设计的一方面，而“设计结合自然”的科学思维要求综合叠加各项因子（如竖向、土壤、径流和植被等）以论证设计决策

的合理性。对水文水动力模型和物联网技术等物理模型和监测技术的应用[29]~[31]也有助于增加设计的可解释性。因而，如何连接GAN模型代表的形态表达与物理模型代表的定量分析是AI深度融入设计学科必然要克服的问题。华南农业大学的研究团队已经尝试耦合GAN生成方法和循证健康设计策略来满足适老花园设计的目标[32][33]。随着GAN生成布局多样性的提升，未来利用多目标优化算法对其进行筛选、优化将有助于提升设计决策的科学性[34]。而随着生成算法的更新，物理模型及优化算法将有可能逐步与AI模型融合，显著提升GAN生成方法的可解释性和应用深度。

致谢

感谢参与问卷调查的湖南大学建筑与规划学院、清华大学建筑学院的师生及北京市市政工程设计研究总院的各位同仁。感谢上海交通大学易冉助理教授及清华大学刘海龙长聘副教授对本研究的帮助与支持。

补充材料

可通过https://doi.org/10.15302/J-LAF-1-020094查看本文补充材料。