

AIGC-Empowered Logic-Driven Generative Framework for Chinese Classical Gardens: Bridging Cultural Semantics and Spatial Constraints

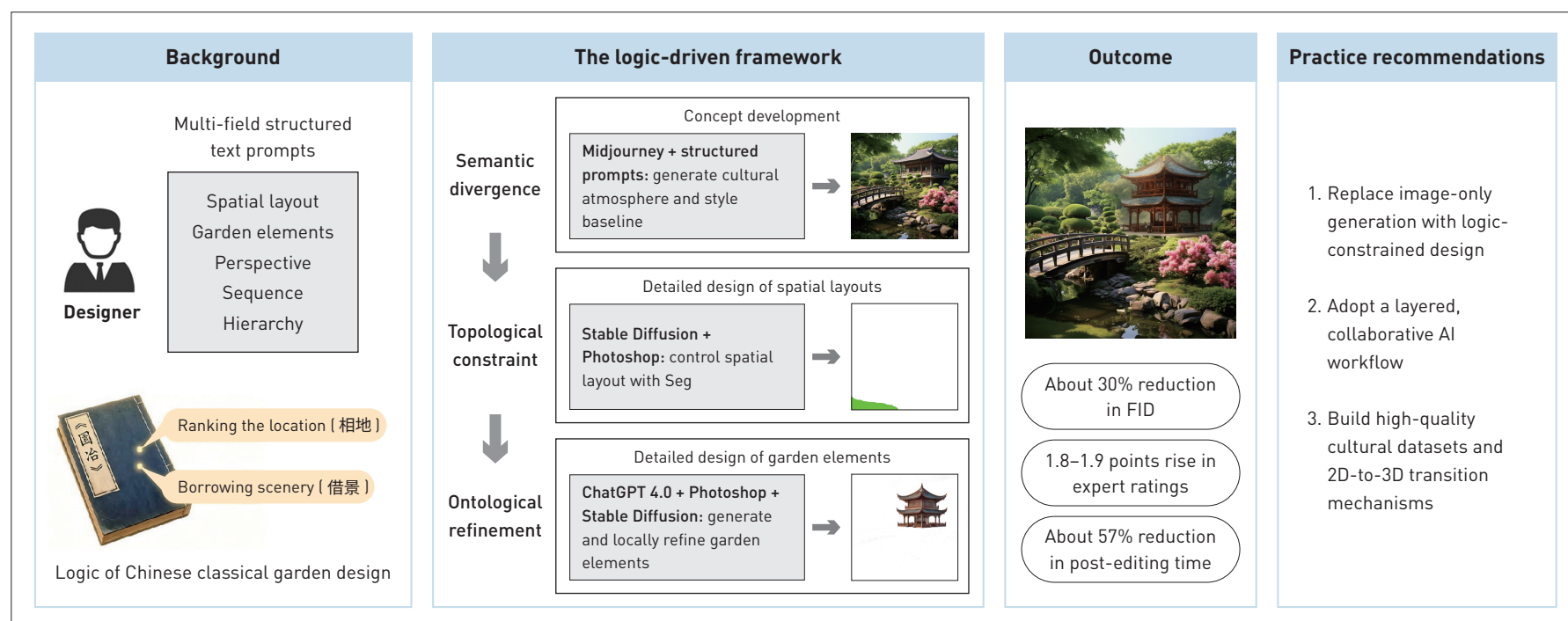
Chenfeng XU^{1,2}, Huilin LIANG^{1,*}

¹ College of Landscape Architecture, Nanjing Forestry University, Nanjing 210037, China
² School of Architecture, Tianjin University, Tianjin 300072, China

*CORRESPONDING AUTHOR

Address: No. 159, Longpan Road, Xuanwu District, Nanjing 210037, China
Email: huilinliang@njfu.edu.cn

GRAPHICAL ABSTRACT



ABSTRACT

Recent advances in artificial intelligence have introduced new opportunities for landscape architecture, yet high-context domains such as Chinese classical garden design remain challenged by dataset bias and spatial hallucination in general-purpose models. To overcome the limitations of single-platform approaches to balancing cultural semantics and spatial topology, we propose a logic-driven, multi-platform collaborative framework that encodes traditional design principles as computational constraints. The framework integrates semantic divergence to capture “yijing” (意境)

with topological control to ensure spatial coherence. Performance is evaluated using metrics including Fréchet inception distance (FID), Contrastive Language–Image Pre-training, peak signal-to-noise ratio, and structural similarity index measure, together with expert assessments and ablation experiments. Results indicate that compared with a single-platform text-to-image baseline, the full three-layer workflow achieved about 30% reduction in FID and improved expert-rated cultural fidelity and spatial logic by approximately 1.8–1.9 points, while reducing human post-

editing time by about 57%. Layer-wise ablation shows that spatial improvements are mainly attributable to the Topological Constraint Layer, whereas gains in cultural authenticity are driven by the Ontological Refinement Layer. This study not only establishes a reproducible, model-agnostic workflow for cultural heritage design but also provides a theoretical bridge connecting current 2D generative paradigms with future 3D spatial cognition.

KEYWORDS

Chinese Classical Garden; Artificial Intelligence; Artificial Intelligence Generated Content; Design Framework; Design Workflow; Artificial Intelligence Tool

HIGHLIGHTS

- AI relies on visual composition for design but lacks spatial understanding
- AI has a basic grasp of Chinese classical gardens but struggles with layouts and elements
- An AIGC-assisted workflow integrating spatial layouts and garden elements is proposed
- A logic-driven, multi-platform collaborative AIGC design framework is introduced

RESEARCH FUNDS

- National Natural Science Foundation of China (No. 32371940)
- Humanities and Social Science Fund of Ministry of Education of China (No. 23YJCZH119)

1 Introduction

Chinese classical gardens, represented by Jiangnan private gardens and northern imperial gardens, constitute a distinctive tradition in Chinese landscape design and have exerted broad influence on garden and landscape practices worldwide^[1-2]. Rooted in philosophy, literature, and art, their design emphasizes the ideal of “harmony between heaven and humanity” and encodes cultural meanings through coherent spatial organization

and symbolic details.

Over the past few decades, Chinese classical garden design has undergone a significant transformation from traditional hand drawing to parametric design^[3]. The introduction of digital technologies, such as modeling and rendering software, has improved design accuracy and flexibility while also enhancing visual representation, providing designers with more effective decision-support tools^[4]. However, current workflows have yet to achieve a truly intelligent level, as design processes still rely heavily on the expertise and intuition of designers. Given the inherent complexity of Chinese classical garden design, the field imposes high professional skill requirements, resulting in a substantial entry barrier. Additionally, the design process is often labor-intensive. To produce a visually appealing plan, designers typically undertake time-consuming tasks such as modeling, rendering, and image editing. As a result, considerable effort is devoted to presentation and refinement rather than conceptual design. Meanwhile, with the renewed societal interest in Chinese classical garden design, contemporary projects are often shaped by heterogeneous site contexts, functional demands, cultural interpretations, and stakeholder preferences, which together may require more iterative design coordination and refinement. Consequently, reducing design difficulty, improving design quality and efficiency, and streamlining workflows have become critical challenges for the future development of Chinese classical garden design.

With the rapid advancement of artificial intelligence (AI), the design industry is undergoing a profound transformation toward intelligent workflows^[5]. In this paper, the term “AIGC-assisted workflow” refers specifically to processes that leverage Artificial Intelligence-Generated Content (AIGC), emphasizing the role of AI in content generation and its assistive function in augmenting rather than replacing the designer’s creative process. Unlike traditional design methods that focus primarily on form and function, AIGC-assisted creative tools are reshaping creative workflows by reducing limitations related to scale, scope, and technical learning requirements^[6]. Designers can now use low-dimensional inputs, such as prompts, hand-drawn sketches, and rough models, as control elements to generate detailed high-dimensional design outputs. This approach reduces reliance on repetitive drafting and iterative design^[7-8].

The application of AI in design has become increasingly widespread, influencing design paradigms across various spatial scales, from urban and neighborhood planning to architecture and landscape design^[9]. Its influence now extends into subfields including architectural design^[10-15], interior design^[16], and graphic design^[17].

At present, AIGC applications in design practice can generally be categorized into generative design based on intelligent algorithms and assisted design through intelligent platforms^[18-19]. On the algorithmic side, Generative Adversarial Networks (GANs) and their variants (e.g., Pix2Pix, CycleGAN) have been widely employed in the generation of 2D design proposals and image-to-image translation tasks^[20]. Initially, these algorithms were mainly applied to architectural and urban design. For instance, Jige Steven Quan developed the Urban-GAN system for procedural urban design^[21], while Xueqing Li et al. combined GANs with Generative Genetic Design (GGD) to optimize the spatial layout of pedestrian-friendly urban cooling nodes^[22]. Hybrid models combining autoencoders and GANs have been applied to volumetric building form generation^[23]; and Graph-Constrained GANs have been introduced for modular housing floor plan design^[24]. Simultaneously, researchers have begun to explore the integration of intelligent algorithms into landscape architecture, yielding promising preliminary results. These include automatic rendering of landscape visualizations using the PlantoGraphy system^[25], generation and visualization of landscape layout plans using the Pix2Pix-BicycleGAN workflow^[26]. These techniques have been successfully applied in specific contexts such as pocket parks^[27] and floral border compositions^[28], demonstrating strong adaptability and significant scalability. Overall, intelligent algorithm-based research is broad in scope and technically mature, showing strong potential in complex spatial generation, image interpretation, and multi-objective optimization.

At the platform level, intelligent tools such as CityEngine, UrbanSim, Delve, DEEPUD, XKool, and CityCAD have emerged with capabilities for urban form generation and spatial simulation^[29-30]. Concurrently, text-to-image generation platforms such as Stable Diffusion, Midjourney, and DALL·E enable designers to rapidly generate visual scenarios from natural language prompts, facilitating a combination of rational analysis and intuitive creativity. Representative applications include the systematic assessment of platform-specific potentials and limitations in architectural design^[31]. However, research on intelligent algorithms substantially outpaces research on intelligent platforms in both volume and depth, particularly in architecture and urban design, while applications in landscape architecture remain comparatively limited. Furthermore, regardless of whether the approach is algorithm-based or platform-based, most current methods rely predominantly on 2D top-view plans or related imagery to indirectly represent 3D spatial qualities. This 2D image-centric generation logic constrains the ability to address the spatial integrity, layering, and immersive experience fundamental to complex 3D design. As such, a critical

question remains: can approaches that rely primarily on 2D planar compositions adequately support spatially coherent and immersive 3D design outcomes? This issue warrants further theoretical and methodological investigation.

As a vital component of the design field, Chinese classical gardens are entering a new phase of development through the integration of advanced AI technologies. To enable AI to meaningfully contribute to design practice, it is necessary to establish a shared theoretical foundation bridging computer science and the principles of Chinese classical garden design. Currently, AIGC-assisted design for Chinese classical gardens primarily relies on algorithms (e.g., CycleGAN, Pix2Pix, Pix2PixHD) to achieve intelligent and efficient design generation. For example, previous studies have used ANN and GAN models to learn 3D spatial structures and generate Taihu stones^[32] and trained GAN models to generate layout plans for private Chinese courtyards^[33]. However, these methods depend heavily on large-scale, high-quality image datasets and generally require comprehensive feature extraction to produce high-quality outputs. In addition, they often require advanced programming skills, limiting their accessibility in practical design applications. In contrast, the rapid development of widely accessible AIGC platforms has encouraged designers to adopt such platforms as primary creative tools. Recent reviews of AIGC-assisted design and human-AI interaction highlight a paradigm shift from automation toward co-creation^[34-36]. However, as noted in previous studies, most current frameworks focus on modern urban scenarios, while research explicitly addressing the high-context cultural semantics embedded in heritage domains such as Chinese classical gardens remains limited. Therefore, there is a need to develop practical and theoretically grounded workflows based on widely used AIGC platforms.

Critically, this study aims to move beyond the simple aggregation of commercial AI tools by establishing a logic-driven generative framework intended to bridge the gap between high-context cultural semantics and low-context generative AI. The study argues that standard generative models may suffer from data bias and spatial hallucinations when interpreting non-Western architectural forms. Therefore, beyond improving efficiency, the primary contribution of this research lies in the proposed logic-driven generative framework, which addresses three limitations of current AIGC systems. First, unlike conventional text-to-image generation approaches, the framework introduces a Semantic Divergence Layer that functions as a cultural schema retrieval mechanism. By translating abstract classical theories from *Yuan Ye* (《园冶》) into structured prompt logic, this layer aims to reduce semantic drift commonly observed in

general-purpose models. Second, to address the lack of inherent 3D reasoning in current generative systems, the framework proposes a Topological Constraint Layer based on control-guided generation techniques. This layer formalizes the classical principle of ranking the location (相地) as computational edge constraints, compelling the generative process to follow physical spatial logic rather than purely pixel-based probabilities. Third, in response to the limitations of 2D generation, the framework adopts a human-scale sequential viewing strategy through an Ontological Refinement Layer that generates coherent image sequences instead of isolated views. This approach is intended as a transitional methodology bridging current 2D paradigms and future 3D world models, while preserving the logic of “changing scenery with every step” (步移景异) through consistent topological constraints across multiple views. This research seeks to contribute new perspectives to the contemporary development of Chinese classical garden design and to support the continued relevance and accessibility of this cultural heritage in modern society.

2 Methods

2.1 Analyzing the Logic of Chinese Classical Garden Design

The design logic draws directly from classical Chinese aesthetics and philosophical treatises such as Cheng Ji’s *Yuan Ye* ^[37] and Zhenheng Wen’s *Zhang Wu Zhi* (《长物志》) ^[38]. Specifically, the concept of borrowing scenery (借景) guides the Semantic Divergence Layer, allowing the AI to expand visual boundaries while maintaining logical coherence. Similarly, the classical principle of ranking the location from *Yuan Ye* was applied as the theoretical basis for Topological Constraint Layer, treating physical site conditions as rigid computational constraints.

Based on the integrated relationship between spatial layouts and garden elements in Chinese classical garden design^①, this research examined landscape characteristics and design logic from these two complementary perspectives: the organization of spatial layouts and the composition of garden elements. The analysis of garden elements concentrates on four principal components: rockeries, water, vegetation, and architecture.

① The spatial arrangement and combination of Chinese classical garden elements generate complex spatial and temporal relationships, including viewing and being viewed, primary and secondary elements with focal points, spatial contrasts, concealment and exposure, guidance and suggestion, sparsity and density, undulation and layering, void and solid, meandering paths, varied terrain, upward and downward perspectives, penetration and layering, and spatial sequences.

2.2 Analyzing the Ability of Existing AIGC Platforms in Chinese Classical Garden Design

We tested the capabilities of mainstream AIGC platforms, including ChatGPT 4.0, Stable Diffusion, and Midjourney, in layout and element generation to evaluate their understanding of Chinese classical gardens.

1) Viewpoint. Since AIGC platforms primarily generate 2D images, we analyzed their capability to interpret various viewpoints (e.g., plan, elevation, human perspectives, bird’s-eye views). Viewpoint alignment was judged independently by three trained annotators based on horizon line position, camera elevation angle, and consistency between described and rendered occlusion relationships.

2) Spatial layouts and garden elements. The test of spatial layouts and garden elements consisted of preliminary and detailed stages. Preliminary testing utilized approximately 10 general spatial prompts to evaluate baseline comprehension and identify the most capable platforms. Subsequently, detailed testing evaluated these top-performing platforms on their understanding of specific categories within spatial layouts and garden elements.

3) Prompt standardization procedure. To ensure a fair and scientifically rigorous comparison across AIGC platforms with different syntax requirements, we established a systematic prompt standardization procedure. We designed a structured prompt template that decomposes each design concept into five core informational fields: core concept, main elements, spatial layout, art style or atmosphere, and technical parameters.

This core semantic information was then translated into platform-optimal formats: coherent narrative paragraphs for DALL-E 3, and comma-separated keywords for Midjourney and Stable Diffusion (augmented with negative prompts for the latter). This structured approach ensures that the design brief remains consistent, allowing for a valid comparison of their generative performance.

2.3 Multidimensional Evaluation Protocol

To move beyond surface-level image metrics, we established a comprehensive evaluation framework comprising four distinct stages.

1) Technical baseline. We employed established metrics from the field of computer vision and generative model assessment. The Fréchet inception distance (FID) measures the distance between the feature distributions of real and generated images. This research first constructed a reference dataset of over 1,000 high-resolution photographs of real Chinese classical gardens. Then, a corresponding set of 1,000 images from each AIGC platform was generated using a

standardized set of prompts. The FID score was computed between each generated set and the reference set. This research employed the Contrastive Language-Image Pre-training (CLIP) score, which calculates the cosine similarity between the embedding of a text prompt and a generated image. To quantify the impact of specific adjustments to the spatial layout, we utilized the peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM). We calculated the PSNR and SSIM between the pre-refinement and post-refinement images. For the testing of garden elements, approximately 30 preliminary tests were conducted. For each category, 20 detailed tests were conducted.

2) Human-centered user experience. We recruited 72 participants (37 graduate students in Landscape Architecture and 35 junior designers with 1–3 years of practice experience). Each participant was given a standardized 2-hour design task: to develop a conceptual design for a small courtyard corner in the style of a Jiangnan private garden. Each participant completed the System Usability Scale (SUS) questionnaire immediately after the task. In this study, we adopted an extended SUS structure (20 items, each rated on a 5-point Likert scale).

3) Expert cultural turing test. We convened a panel of 26 senior landscape architecture academics and practitioners. A curated set of 20 anonymized images was presented to the panel. The experts were asked to evaluate each image using a 5-point Likert scale across five criteria: cultural fidelity, spatial logic, implementability, yijing (意境)^②, and aesthetic atmosphere. Fleiss Kappa was then calculated. A Kappa value exceeding 0.6 is generally interpreted as indicating substantial agreement.

4) Layer-wise ablation analysis. To attribute observed improvements to specific components rather than tool choice or operator skill, we conducted controlled ablation experiments by treating the Semantic Divergence, Topological Constraint, and Ontological Refinement Layers as independent functional modules. Four ablation variants were compared under identical design tasks and operational conditions (Table 1). Minimal human selection was allowed, while cross-layer compensatory operations were strictly prohibited. By jointly analyzing quantitative metrics (FID, CLIP, PSNR, SSIM), user evaluations, and expert assessments, this comparative analysis enabled explicit identification of each layer’s contribution to spatial coherence, cultural accuracy, and human post-editing effort.

② “Yijing” refers to “意境” in Chinese, which might be translated as “imagery” in relevant literatures.

Table 1: Ablation settings with different combinations of logic layers

Group	Semantic Divergence Layer	Topological Constraint Layer	Ontological Refinement Layer	Description
G0	×	×	×	Single-platform direct generation (Midjourney baseline)
G1	✓	×	×	Semantic divergence only
G2	✓	✓	×	Semantic divergence + topological constraint
G3	✓	✓	✓	Full three-layer logic-driven framework (proposed method)

2.4 Exploring a Generative Framework and Operational Workflow for Chinese Classical Garden Design

In the design process, designers must provide AI tools with comprehensive and accurate prompts that encapsulate project requirements, site characteristics, and professional insights. Currently, guiding AI platforms in the creative process is an inherently iterative process of continuously providing information to the AI tools and refining concepts and ideas based on feedback. In this process, the quality and structure of these prompts directly dictate the efficiency and quality of the generated outputs.

The fundamental logic of Chinese classical garden design relies on integrating “spatial layouts” and “garden elements” to construct layered 3D spaces. However, current AIGC platforms lack inherent 3D spatial cognition and struggle to accurately interpret 3D characteristics of landscape spaces from 2D plan views. To bridge this gap, we proposed using 2D landscape images from non-plan perspectives as the primary design medium, dividing the workflow into two distinct stages: conceptual design and detailed design.

1) Conceptual design. Utilizing platforms with strong text-to-image capabilities, designers input detailed prompts to generate initial concepts and stylistic directions. To ensure rigorous evaluation, a conceptual design iteration is considered acceptable only when all major spatial layouts, garden elements, and stylistic markers described in the prompt are present without structural contradictions.

2) Detailed design. Building upon the accepted conceptual images, this stage involves precise, localized adjustments to resolve spatial issues such as disproportionate layouts, unclear

path organization, and disconnected water-land relationships. A modification during this phase is quantitatively and qualitatively defined as successful when the PSNR or SSIM values match the intended magnitude of change, and human annotators confirm that the spatial logic is demonstrably improved relative to the previous iteration.

Finally, this optimized logic-driven workflow was applied to a controlled case study of a Jiangnan private garden, covering an area of approximately 500 m² and characterized by small bridges and flowing water, to explore the practical application of AIGC in Chinese classical garden design.

3 Results

3.1 The Logic of Chinese Classical Garden Design

Through a systematic analysis of Chinese classical gardens, we first abstracted the complex design logic of Chinese classical gardens into two computationally evaluable dimensions, i.e., spatial layouts and garden elements, and summarized various garden elements, including rockeries, water, vegetation, and architecture. This analysis provides a theoretical basis for evaluating the AI’s ability to interpret and generate Chinese classical gardens.

3.1.1 Spatial Layouts

Chinese classical gardens emphasize the concept of “grandness within smallness” (小中见大), skillfully utilizing limited space to create diverse and rich scenic effects. By guiding visitors’ sight and walking paths, the design enhances the overall experience. The spatial layouts of Chinese classical gardens generally fall into two types: dominant scene layout and clustered scene layout. The dominant scene layout uses techniques such as elevation changes and axial symmetry to highlight the main scenery, enhancing its prominence and visual appeal. In contrast, the clustered scene layout seemingly arranges elements freely and loosely but actually maintains a strong sense of order to prevent the scattered elements from seeming chaotic. The techniques employed in spatial layouts include borrowed scenery (借景), framed scenery (框景), leaking scenery (漏景), opposing scenery (对景), layered scenery (夹景), obstructed scenery (障景), added scenery (添景), and hidden scenery (藏景)^[39].

3.1.2 Garden Elements

(1) Rockeries

In Chinese classical gardens, rockeries are constructed using natural stones^[40] to emulate the grandeur of natural mountains.

The selection of stones emphasizes their natural beauty, including Taihu stone (太湖石), lingbi stone (灵璧石), ying stone (英石), and kun stone (昆山石)^[41]. The basic techniques for constructing rockeries can be categorized into various methods, such as an (安), lian (连), jie (接), dou (斗), pin (拼), kua (挎), xuan (悬), jian (剑), ka (卡), chui (垂), tiao (挑), cheng (撑), die (叠), shu (竖), jia (夹), zheng (整), and gou (钩). The surface texture treatment techniques of rockeries, known as cunfa (皴法), include various styles such as mayacun (马牙皴), pimacun (披麻皴), fupicun (斧劈皴), zhedaicun (折带皴), and luanchacun (乱柴皴).

(2) Water

In the creation of waterscapes, two primary layout forms are employed: centralized layouts and dispersed layouts. The design aims to achieve various artistic effects, such as graceful water curves, the natural rhythm of water sounds, the dynamic and static interplay of water bodies, and the enchanting beauty of reflections on the water surface^[42]. Water is typically integrated closely with the spatial layouts. Through its interaction with other garden elements, water creates diverse spatial variations, significantly enhancing the vibrancy and layering of the entire garden.

(3) Vegetation

Chinese classical gardens place great emphasis on their multifaceted appreciation of vegetation, including their color, fragrance, shape, form, and charm. Planting methods primarily adopt natural forms and various types such as paired planting (对植), row planting (列植), clustered planting (丛植), and solitary planting (孤植). The main categories of plants include trees, flowers, bamboo, and herbaceous plants^[43]. Different plants are often imbued with symbolic meanings, such as elegance, resilience, and prosperity.

(4) Architecture

The landscape composition of Chinese classical gardens is often realized through a variety of architectural forms or structures. The architectural types within gardens are diverse, including pavilions (亭), terraces (台), towers (楼), chambers (阁), verandas (轩), open halls (榭), archways (卷), open spaces (广), and corridors (廊)^[44]. Depending on the specific setting and functional requirements, the roof styles also vary, with common types including wudian (庑殿), xieshan (歇山), xuanshan (悬山), yingshan (硬山), and cuanjian (攒尖).

3.2 Objective and Subjective Evaluation Results

3.2.1 Technical Baseline

The generation performance of the three AIGC platforms was initially benchmarked using FID, CLIP, PSNR, and SSIM metrics (Table 2). Midjourney achieved the lowest FID (8.4) and the highest

Table 2: Quantitative evaluation of AIGC platforms

Platform	FID score	Average CLIP score	Average PSNR score	Average SSIM score
Midjourney	8.4	0.2	42.9	0.9
Stable Diffusion	11.3	0.3	36.1	0.7
DALL-E 3	28.7	0.5	37.7	0.8

visual quality scores compared with Stable Diffusion and DALL-E 3, indicating superior baseline generation capabilities.

3.2.2 Human-Centered Evaluation

Furthermore, a usability study with 72 participants evaluating the multi-platform workflow yielded a mean SUS score of 76.6, reflecting a highly positive user experience, particularly in efficiency enhancement and iterative design support. The results also indicate ongoing limitations of current image-based AIGC systems in representing complex spatial perception and experiential continuity.

3.2.3 Expert Cultural Turing Test

While Midjourney achieved the best baseline technical scores, expert evaluation (Fleiss Kappa = 0.8) revealed a critical divergence between visual appeal and design logic. As shown in Fig. 1, although single-platform outputs achieved high scores in aesthetic atmosphere, they performed significantly worse in spatial logic due to structural hallucinations, such as physically implausible bridge structures. In particular, the generated bridges often lacked proper integration with the path system, resulting in low implementability scores. In contrast, our proposed multi-platform framework achieved a balanced high score across all dimensions, particularly outperforming single-platform baselines in cultural fidelity (4.5/5.0) and spatial logic (4.2/5.0).

3.2.4 Layer-wise Ablation Analysis

To systematically attribute the observed improvements to specific components of the framework, a layer-wise ablation analysis was conducted. Compared with the G0 baseline, the full three-layer workflow (G3) achieved consistent improvements across generation quality, expert evaluation, and editing efficiency. FID decreased from 68.42 ± 4.21 to 47.92 ± 2.93 , corresponding to a 30.0% reduction, while SSIM and PSNR increased from 0.62 ± 0.05

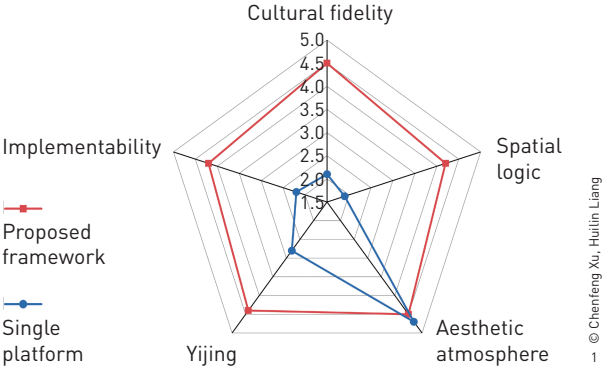


Fig. 1
Comparative radar chart.
Fig. 2
Comparison of generation effects of different platforms.

to 0.73 ± 0.03 and from 17.31 ± 1.10 to 21.15 ± 0.81 , respectively. Expert-rated cultural fidelity improved by 1.92 (2.63 to 4.55), and spatial logic improved by 1.82 (2.42 to 4.24). Meanwhile, human post-editing time was reduced from 42.02 ± 8.51 to 18.01 ± 5.26 min, a 57.1% decrease. These results suggest that G3 substantially enhances cultural-spatial consistency while reducing manual correction effort.

3.3 A Comprehensive Workflow for Chinese Classical Garden Design

The testing results show that while images generated by DALL-E 3 and Stable Diffusion generally align with the input prompts, their visual quality remains relatively average. In contrast, Midjourney not only achieves a high degree of prompt adherence but also produces visually superior results (Fig. 2). Furthermore, the findings indicate that an iterative, three-stage input strategy—defining task objectives, structuring a detailed prompt, and optimizing based on outputs—substantially improves the consistency and visual fidelity of AI-generated images (Fig. 3).

Our experiments show that an iterative input process comprising four stages—“prepare accurate spatial layout color block diagrams,” “modify error spatial layouts,” “input positive and negative prompts,” and “adjust and optimize”—can significantly enhance the stability and aesthetic quality of AI-generated images and yield more precise design outputs and clearer spatial layout representations.



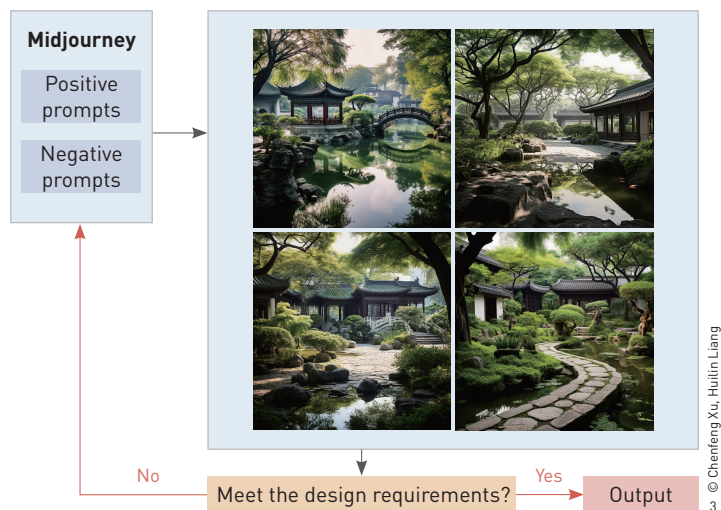


Fig. 3
Operational
workflow of
conceptual
design.

Evaluations (Table 3) indicate that Stable Diffusion demonstrates exceptional control and generation performance in the “Seg” type, accurately mapping different color blocks to corresponding landscape elements with a high degree of alignment. This precise spatial control is primarily achieved through ControlNet, a neural network architecture designed to inject extra spatial conditioning

Table 3: Effect evaluation of different ControlNet types

Type	Characteristics	Advantage	Disadvantage
Canny	Hard edge extraction	Smoothed and not affected by the gray value of the image	Rough treatment and without thick depth
Soft edge	Soft edge extraction	Soft lines	Fuzzy edge of the extraction line
Lineart	Line drawing extraction	Rich in detail, diverse processing types	Unstable results, frequent errors in understanding complex scenes
Scribble	Doodle/sketch extraction	Greater freedom for imagination	Poor detail capture of garden elements, frequent confusion between water bodies and ground
Seg	Semantic segmentation extraction	Precise identification of spatial layouts and garden elements	Low accuracy for garden elements
Others	Not suitable for Chinese classical garden design		

into pre-trained diffusion models. The remaining types are not well suited to Chinese classical garden design.

The underlying mechanism of ControlNet is highly effective for architectural applications. Instead of retraining the entire foundational model, it creates a locked copy of the production-ready Stable Diffusion model to preserve its vast visual knowledge, while a trainable copy learns the relationship between an input control map (e.g., Canny edge map, semantic segmentation map) and the desired output. During generation, these outputs of both the locked and trainable copies are merged, guiding the final image via both semantic text prompts and precise spatial constraints. This architecture is uniquely suited for the detailed design phase, as it imposes explicit structural and layout information onto a generated concept without degrading the rich visual knowledge embedded in the base model. Therefore, optimizing reference images from the conceptual design phase by applying localized repainting with the “Seg” tool can effectively enhance the controllability and quality of the generated design outputs.

Within this semantic segmentation approach, utilizing the ADE20K dataset enables Stable Diffusion to generate corresponding elements with high precision based on predefined colors. Further analysis reveals that reclassifying and integrating the four major categories of garden elements improve both the scientific rigor and the practical applicability of AI-generated images in landscape design.

The test results indicate that ChatGPT 4.0 (DALL-E 3) demonstrates outstanding performance in understanding and representing garden elements, generating more precise design images through well-structured prompt combinations (Fig. 4-1). Further analysis reveals that an iterative input process comprising three stages—“assign a role,” “define garden element attributes,” and “adjust and optimize”—can significantly enhance the stability and aesthetic quality of the generated images. The specific tasks for this three-stage iterative process for garden element generation are detailed in the supplementary information.

Finally, test results indicate that current AIGC platforms remain incapable of seamlessly generating complex spatial layouts and precise garden elements simultaneously, necessitating their decoupled processing. This step plays a critical role in optimizing the size, positional control, and accuracy of garden elements within the spatial layout (Fig. 4-2). Integrating the “Canny” tool with conventional image editing software (e.g., Photoshop) ensures precise visual refinement. Adjusting parameters through a three-stage process—“prepare accurate garden elements,” “remove error garden elements,” and “adjust and optimize”—can significantly improve the stability and aesthetic quality of the generated images.

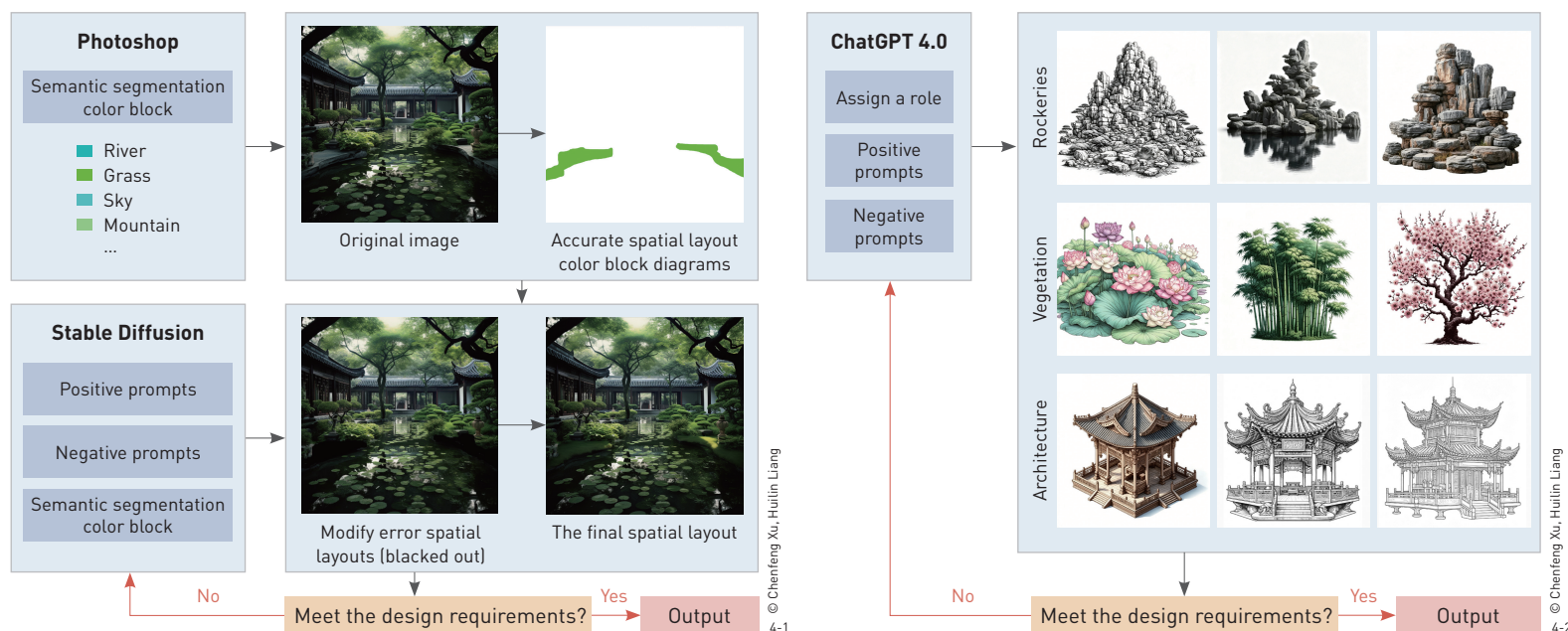


Fig. 4 Detailed design of spatial layout and garden elements.

Fig. 5 The proposed framework (as long as the two underlying logics can be implemented, the choice of platform is not restricted).

Based on the theoretical framework established in Section 2.3 and the test results mentioned above, we finalized the operational workflow as shown in Fig. 5. The validation results confirm that separating the process into conceptual design and detailed design effectively overcomes the AI's limitations in one-shot generation. In the detailed design stage, the workflow explicitly decouples the optimization of spatial layouts and garden elements to ensure distinct control over topological logic and semantic fidelity.

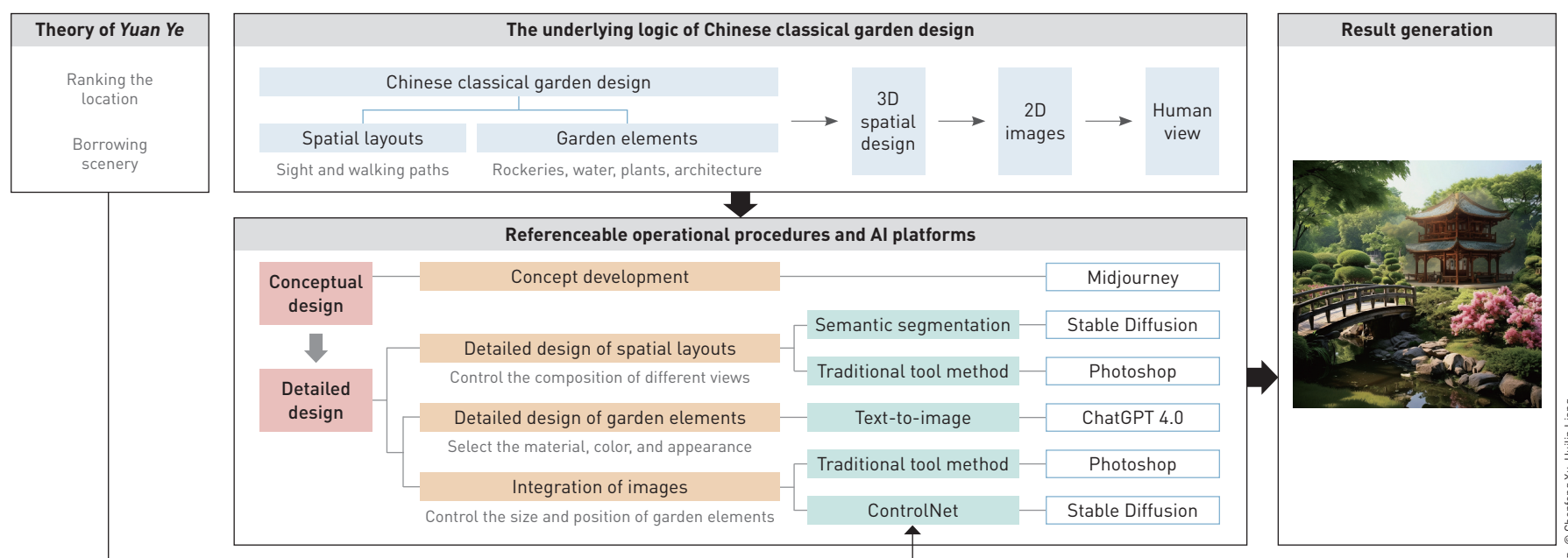
3.4 Case Study

To demonstrate the scientific rigor and reproducibility of our

framework, we applied it to a controlled design experiment: a 500 m² Jiangnan private garden. Unlike utilizing a black-box generation process, we documented the specific AI errors encountered at each stage and how our layered framework corrected them (Fig. 6).

1) Stage 1: Initial Semantic Divergence Layer using Midjourney provided a strong atmospheric baseline. However, structural analysis revealed significant spatial errors, specifically a physically impossible bridge structure that disconnected the water system. This confirmed that pure semantic generation lacks topological logic.

2) Stage 2: To correct the spatial error, we did not rely on



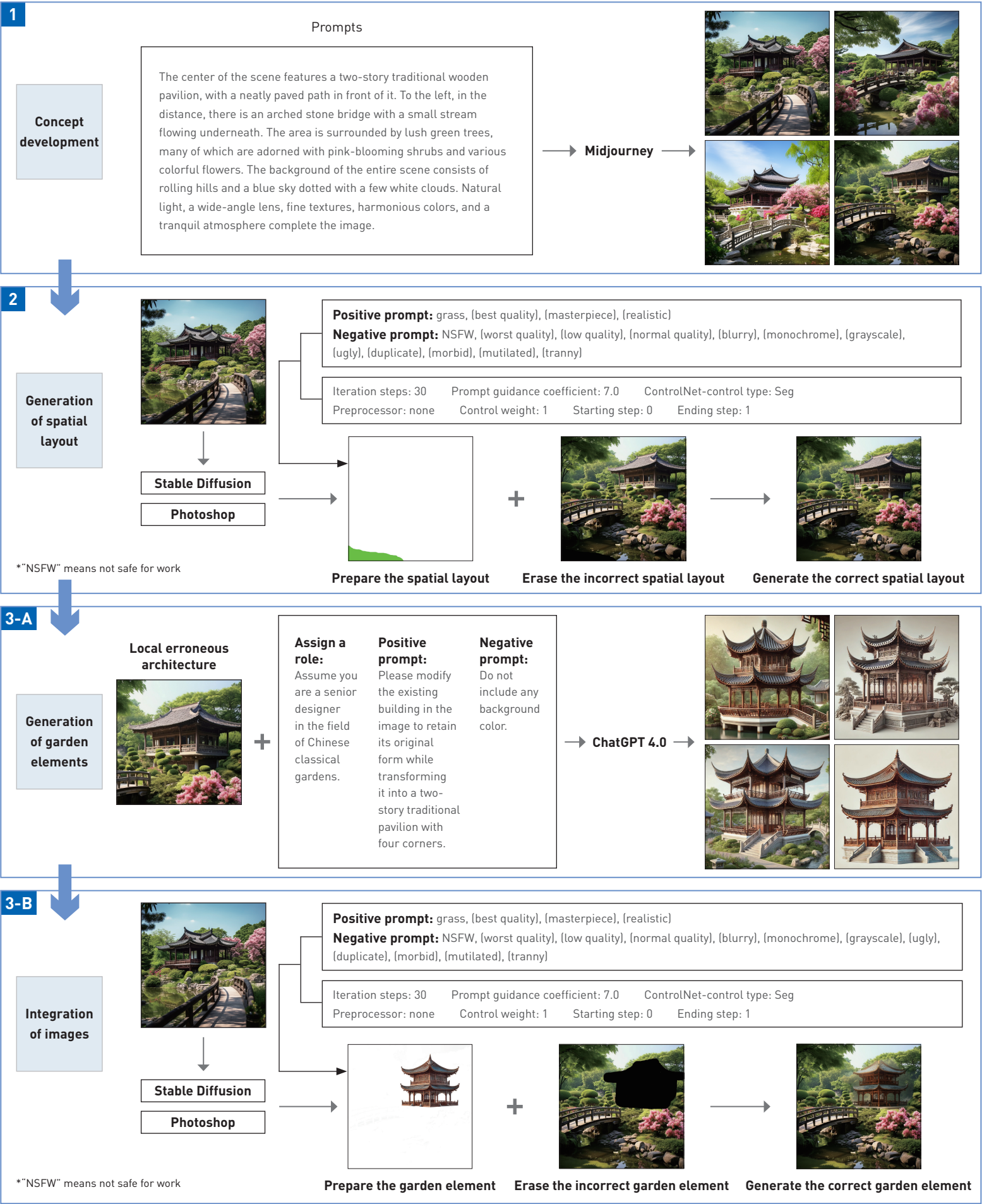


Fig. 6
Process of
the case
study.

random regeneration. Instead, Topological Constraint Layer was applied. By extracting the Seg map of a functionally correct layout and injecting it via ControlNet (weight 1), we forced the model to reconstruct the bridge with correct perspective and connection logic, effectively imposing physical laws onto the generative process.

3) Stage 3-A: The initial AI output rendered the pavilion with generic Asian-style roofs, exhibiting Semantic Drift. We employed the Ontological Refinement Layer using localized inpainting strategies. By constraining the generation to specific architectural vocabulary, we restored the cultural authenticity of the artifacts.

4) Stage 3-B: The final output was not merely selected for aesthetics but validated against the established expert evaluation criteria. The rigorous correction process demonstrates the reproducibility of the workflow: by constraining topology and refining ontology, designers can consistently achieve culturally accurate outcomes, independent of random seed variations.

4 Discussion

4.1 The Adaptability of the Framework in the Era of Rapid AI Evolution

The field of generative AI is evolving at a breakneck pace, with newer models such as Google’s Gemini 2.5 and Black Forest Labs’

FLUX.1 offering superior prompt adherence and integrated editing capabilities. However, the emergence of these advanced models does not render the proposed workflow obsolete; rather, it highlights the resilience of the logic-driven generative framework proposed in this study. The framework is designed as a modular system where specific tools function as interchangeable components within three logic layers (Table 4).

1) Semantic Divergence Layer. Currently powered by Midjourney, this layer can be seamlessly upgraded to newer engines like FLUX.1. While new models improve textual understanding, the methodological necessity of this layer—to retrieve cultural schemas before defining structure—remains unchanged.

2) Topological Constraint Layer. While models like Gemini 2.5 offer editing, they often lack the granular, pixel-level structural control provided by ControlNet. For Chinese classical gardens, where spatial relationships must follow strict physical logic (ranking the location), the black-box editing of end-to-end models is often insufficient. The framework’s insistence on explicit topological constraints remains a critical theoretical contribution.

3) Ontological Refinement Layer. General-purpose models still suffer from dataset bias regarding high-context cultural heritage. The potential integration of LoRA fine-tuning represents a lasting strategy for ensuring cultural fidelity, a requirement that persists regardless of the base model’s power.

Table 4: Correspondence between design logic layers and AI technical architectures

Logic layer	Classical principle	Core technical architecture	Representative model	Role in framework
L1: Semantic Divergence	Borrowing scenery	Proprietary diffusion/flow transformers prioritizing aesthetics and prompt adherence	Midjourney v6 FLUX.1	Cultural schema retrieval: generating atmospheric baselines and overcoming blank canvas paralysis by retrieving high-dimensional cultural features
L2: Topological Constraint	Ranking the location	Control-guided latent diffusion injecting explicit spatial conditions via edge/depth maps	Stable Diffusion + ControlNet (Seg/Depth)	Spatial structure enforcement: forcing the generative process to adhere to physical site logic, such as gravity and spatial connectivity, thereby preventing spatial hallucinations
L3: Ontological Refinement	Physiognomy of things	Fine-tuned diffusion weights (e.g., low-rank adaptation [LoRA]) or precision inpainting	Stable Diffusion (support for LoRA/fine-tuning)	Cultural fidelity restoration: correcting specific heritage forms, such as Taihu stones and flying eaves, that are often misinterpreted by general-purpose models due to dataset bias through localized refinement or custom weighting strategies

4.2 Capability Boundaries and the Path to World Models

To assess the long-term stability of the proposed workflow, it is necessary to situate it within the fundamental principles of current AI pathways. Mainstream models, such as Stable Diffusion and Midjourney, are based on latent diffusion and rely on probabilistic distributions to map textual inputs to 2D pixels. Although effective in rendering texture and stylistic features, these models lack an internal representation of physical laws, which often results in spatial inconsistencies and geometric errors. Emerging advances in 3D generative models, including NeRF and 3D Gaussian Splatting, as well as world models such as Sora and Gen-3, aim to address these limitations by incorporating physical simulation and object permanence. However, currently, these 3D models often sacrifice texture fidelity and cultural specificity for geometric accuracy. In this transitional era, the human-perspective-first strategy proposed in this study serves as a critical bridge. By generating spatially consistent 2D image sequences via the Topological Constraint Layer, the workflow effectively acts as a semantic frontend. These culturally accurate 2D sequences can serve as high-quality input for subsequent 3D reconstruction algorithms, such as image-to-3D NeRF, ensuring that the cultural essence established by the framework is retained as the spatial representation transitions to fully generated 3D forms. Consequently, the core logic of the framework remains applicable as underlying generative engines evolve from 2D to 3D paradigms.

4.3 Addressing Cultural Homogenization in Heritage Design

While established commercial tools, such as Photoshop (Firefly) and FLUX.1 Inpainting, offer streamlined workflows for image composition and refinement, applying these general-purpose tools directly to Chinese classical garden design reveals a critical limitation: cultural homogenization. Commercial black-box models are typically trained on copyright-cleared, globalized datasets dominated by Western or modern architectural forms. Comparative testing indicates that when these tools are applied to the restoration of specific heritage elements, such as Taihu stones, they tend to produce generic rock forms that lack the defining aesthetic qualities of shou (瘦), zhou (皱), lou (漏), and tou (透). Similarly, they frequently misinterpret the complex curvature of flying Eaves (翘角) as standard pagoda roofs. In contrast, the innovation of the proposed workflow lies not in the act of inpainting itself, but in the modeling of cultural semantics through a controllable, open ecosystem. By utilizing the open Stable Diffusion ecosystem, the framework allows for the potential injection of specific cultural datasets via LoRA fine-tuning strategies within the

Ontological Refinement Layer. Although the current study validates this layer using iterative inpainting, the framework provides the necessary technical infrastructure for future implementation of custom cultural models, ensuring that spatial sequencing and element generation are not merely visually plausible but culturally accurate. Therefore, this methodology transcends the convenience of commercial tools, offering a necessary technical pathway for preserving cultural sovereignty in the age of generative AI.

4.4 Research Objectives and Tool Limitations

It is crucial to distinguish between the inherent limitations of current generative tools and the methodological objectives of this study. While current paradigms face constraints regarding innate 3D understanding and full automation, the primary objective of this research is not to create a fully automated 3D modeling tool, but to establish a logic-driven methodology for translating the high-context semantics of Chinese classical garden into computational constraints.

5 Conclusions

The relationship between designers and AI is shifting from a simple user-tool dynamic to one of deep collaboration and co-creation. To achieve this ideal state, our research has adopted several key approaches: conducting an in-depth analysis of the characteristics and inherent logic of Chinese classical garden design, evaluating the capabilities of mainstream AI platforms, and integrating the strengths of these platforms to develop an innovative and efficient AIGC-assisted workflow for Chinese classical garden design.

This research has the following results. Firstly, Chinese classical gardens possess unique characteristics with their inherent logic centered on the organization and planning of spatial layouts and garden elements. In the context of current AI generation, this process often involves the complex task of translating 3D spaces into 2D representations. Secondly, different mainstream AI platforms demonstrate distinct advantages: Midjourney excels in providing design inspiration, ChatGPT 4.0 (DALL-E 3) specializes in generating accurate garden elements, and Stable Diffusion demonstrates superior performance in controlling and optimizing spatial layouts. Thirdly, while AI has shown preliminary design capabilities, achieving high-quality outcomes still requires the integration of traditional tools and professional input from designers. In the future, strengthening training datasets and optimizing technologies can better support the intelligent

transformation of Chinese classical garden design.

This research represents an initial foundational framework of integrating AI with Chinese classical garden design, marking a new starting point for this field. Through this exploratory work, we have not only opened up new directions for the digitalization and intelligent development of Chinese classical garden design but also laid a solid foundation for future in-depth studies in this area. Our findings provide scholars with a scientific reference framework, encouraging further innovative research and driving the transformation and enhancement of Chinese classical garden design in the context of modern technology.

ELECTRONIC SUPPLEMENTARY MATERIAL

Supplementary material is available in the online version of this article at <https://doi.org/10.15302/J-LAF-2026-0033>.

REFERENCES

- [1] Tran, D. K., Ignatieva, M., & Tenorio, R. (2023). Exploring classical Vietnamese gardens under the Nguyen Dynasty: A comparative study with Chinese counterparts. *Landscape Architecture Frontiers*, 11(5), 44–58.
- [2] Yang, H. (2022). *A Treatise on the Garden of Jiangnan—A Study on the Art of Chinese Classical Garden*. Springer.
- [3] Guo, L., Ma, W., Gong, X., Zhang, D., Zhai, Z., & Li, M. (2024). Digital preservation of classical gardens at the San Su Shrine. *Heritage Science*, 12, 66.
- [4] Shan, P., & Sun, W. (2021). Research on landscape design system based on 3D virtual reality and image processing technology. *Ecological Informatics*, 63, 101287.
- [5] McGuire, J., De Cremer, D., & Van de Cruys, T. (2024). Establishing the importance of co-creation and self-efficacy in creative collaboration with artificial intelligence. *Scientific Reports*, 14, 18525.
- [6] Huang, Y.-C., Wensveen, S., & Funk, M. (2023). Experiential speculation in vision-based AI design education: Designing conventional and progressive AI futures. *International Journal of Design*, 17(2), 1–17.
- [7] Albukhari, I. N. (2025). The role of artificial intelligence (AI) in architectural design: A systematic review of emerging technologies and applications. *Journal of Umm Al-Qura University for Engineering and Architecture*, 16, 1457–1476.
- [8] Tan, L., & Luhrs, M. (2024). Using generative AI Midjourney to enhance divergent and convergent thinking in an architect's creative design process. *The Design Journal*, 27(4), 677–699.
- [9] Jiang, F., Ma, J., Webster, C. J., Chiaradia, A. J. F., Zhou, Y., Zhao, Z., & Zhang, X. (2024). Generative urban design: A systematic review on problem formulation, design generation, and decision-making. *Progress in Planning*, 180, 100795.
- [10] Bao, Y., & Xiang, C. (2024). Exploration of conceptual design generation based on the deep learning model—Discussing the application of AI generator to the preliminary architectural design process. *Creativity in the Age of Digital Reproduction*, 343, 171–178.
- [11] Paananen, V., Oppenlaender, J., & Visuri, A. (2023). Using text-to-image generation for architectural design ideation. *International Journal of Architectural Computing*, 22(3), 458–474.
- [12] Chen, J., Wang, D., Shao, Z., Zhang, X., Ruan, M., Li, H., & Li, J. (2023). Using artificial intelligence to generate master-quality architectural designs from text descriptions. *Buildings*, 13(9), 2285.
- [13] Alyıldız, E. C. (2023). Generative text-to-image models in architectural design: A study on relationship of language, architectural quality and creativity. *Icontech International Journal of Surveys, Engineering, Technology*, 7(3), 12–26.
- [14] Al-Tabeeb, A. K., & Al-Desouqi, A. (2023). Metaverse in architecture: An approach to documenting and exploring the Egyptian heritage through metaverse. *Green Building & Construction Economics*, 4(2), 276–295.
- [15] Ahmed, H. (2023). The influence of Fatimid architecture heritage on sustainable design of contemporary architectural facades. *International Design Journal*, 13(1), 381–392.
- [16] 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.

- [17] Chen, H.-C., & Chen, Z. (2023). Using ChatGPT and Midjourney to generate Chinese landscape painting of Tang poem ‘The Difficult Road to Shu’. *International Journal of Social Sciences and Artistic Innovations*, 3(2), 1–10.
- [18] Jang, S., Roh, H., & Lee, G. (2025). Generative AI in architectural design: Application, data, and evaluation methods. *Automation in Construction*, 174, 106174.
- [19] Li, C., Zhang, T., Du, X., Zhang, Y., & Xie, H. (2025). Generative AI models for different steps in architectural design: A literature review. *Frontiers of Architectural Research*, 14(3), 759–783.
- [20] Zhou, S., Jia, W., Diao, H., Geng, X., Wu, Y., Wang, M., ... Wu, Z. (2025). A CycleGAN-Pix2pix framework for multi-objective 3D urban morphology optimization: Enhancing thermal performance in high-density areas. *Sustainable Cities and Society*, 126, 106400.
- [21] Quan, S. J. (2022). Urban-GAN: An artificial intelligence-aided computation system for plural urban design. *Environment and Planning B: Urban Analytics and City Science*, 49(9), 2500–2515.
- [22] Li, X., Lu, W., Peng, Z., Zhang, Y., & Huang, J. (2024). Generative design of walkable urban cool spots using a novel heuristic GAN x GAN approach. *Building and Environment*, 266, 112027.
- [23] Zhuang, X., Ju, Y., Yang, A., & Caldas, L. (2023). Synthesis and generation for 3D architecture volume with generative modeling. *International Journal of Architectural Computing*, 21(2), 297–314.
- [24] Liu, J., Qiu, Z., Wang, L., Liu, P., Cheng, G., & Chen, Y. (2024). Intelligent floor plan design of modular high-rise residential building based on graph-constrained generative adversarial networks. *Automation in Construction*, 159, 105264.
- [25] Wang, Y., Xie, L., & Huang, M. (2026). Landscape design concept generation combining cultural mapping technology and multimodal modeling. *Scientific Reports*, 16, 1505.
- [26] Zhou, H., & Xiang, S. (2024). Applicability evaluation and reflection on artificial intelligence-based “image to image” generation of landscape architecture masterplans. *Landscape Architecture Frontiers*, 12(2), 58–67.
- [27] Ren, T., Zeng, Z., Hao, X., & Zhang, X. (2025). A generative optimization framework for pocket parks supported by mature plant carbon sequestration data. *Urban Forestry & Urban Greening*, 112, 128993.
- [28] Feng, L., Sun, Y., Yu, C., Chen, R., & Zhao, J. (2025). The applicability of two generative adversarial networks to generative plantscape design: A comparative study. *Land*, 14(4), 746.
- [29] Luo, Y., He, J., & He, Y. (2017). A rule-based city modeling method for supporting district protective planning. *Sustainable Cities and Society*, 28, 277–286.
- [30] Wang, X., Song, Y., & Tang, P. (2020). Generative urban design using shape grammar and block morphological analysis. *Frontiers of Architectural Research*, 9(4), 914–924.
- [31] Hanafy, N. O. (2023). Artificial intelligence’s effects on design process creativity: “A study on used A.I. text-to-image in architecture”. *Journal of Building Engineering*, 80, 107999.
- [32] Deng, Q., Li, X., Liu, Y., & Hu, K. (2023). Exploration of three-dimensional spatial learning approach based on machine learning—Taking Taihu stone as an example. *Architectural Intelligence*, 2, 5.
- [33] Aalaei, M., Saadi, M., Rahbar, M., & Ekhlassi, A. (2023). Architectural layout generation using a graph-constrained conditional Generative Adversarial Network (GAN). *Automation in Construction*, 155, 105053.
- [34] Li, Y., Chen, H., Yu, P., & Yang, L. (2025). A review of artificial intelligence in enhancing architectural design efficiency. *Applied Sciences*, 15(3), 1476.
- [35] Salma, Z., Hijón-Neira, R., & Pizarro, C. (2025). Designing co-creative systems: Five paradoxes in human–AI collaboration. *Information*, 16(10), 909.
- [36] Yan, S., Wu, C., & Zhang, Y. (2025). Generative design for architectural spatial layouts: A review of technical approaches. *Journal of Asian Architecture and Building Engineering*, 1–21.
- [37] Ji, C. (2009). *Yuan Ye* [The craft of gardens]. Chongqing Publishing House.
- [38] Wen, Z. (2012). *Zhang Wu Zhi* [Treatise on superfluous things]. Zhonghua Book Company.
- [39] Chen, Y., Gu, Y., Liu, Y., & Cao, L. (2024). Unveiling the dynamics of “scenes changing as steps move” in a Chinese classical garden: A case study of Jingxinzhai Garden. *Heritage Science*, 12, 131.
- [40] Siu, V. M. (1999). China and Europe intertwined: A new view of the European sector of the Chang Chun Yuan. *Studies in the History of Gardens & Designed Landscapes*, 19(3–4), 376–393.
- [41] Wang, Z., Song, P., Zhang, Q., Wei, T., & Pan, B. (2024). Digital improvements in the design and construction process of classical Chinese garden rockeries: A study based on material digitization. *Heritage Science*, 12, 327.
- [42] Sun, M., & Bao, Y. (2025). Effects of reflecting water on viewers’ visual attention, spatial perception, and emotional responses: A case study in Chinese heritage garden. *npj Heritage Science*, 13(1), 26.
- [43] Lin, S., & Liu, W. (2024). Cultural connotation and bionic design application of foliage plants in Chinese gardens. *Arts & Communication*, 2(1), 1854.
- [44] Zhu, Y., Lee, S.-H., & Choi, K.-R. (2022). Analysis of the space and design of Chinese, Japanese and Korean academy gardens. *Land*, 11(9), 1474.

AIGC赋能的中国古典园林的逻辑驱动生成框架： 连接文化语义与空间约束

许沉风^{1,2}, 梁慧琳^{1,*}

¹ 南京林业大学风景园林学院, 南京 210037

² 天津大学建筑学院, 天津 300072

*通信作者

地址: 南京市玄武区龙蟠路159号

邮编: 210037

邮箱: huilinliang@njfu.edu.cn

摘要

人工智能的最新进展为风景园林学科带来了新的发展机遇。然而, 在中国古典园林设计这类高语境领域中, 通用模型仍面临数据集偏差与空间幻觉等问题。为克服单一平台方法在平衡文化语义与空间拓扑关系方面的局限, 本文提出了一种逻辑驱动的多平台协同生成框架, 将传统设计原则编码为可计算的约束条件。该框架将用于捕捉“意境”的语义发散机制与保障空间连贯性的拓扑控制方法相结合。研究采用弗雷歇初始距离、对比语言-预训练得分、峰值信噪比、结构相似性指数等指标, 并结合专家打分法与分层消融实验对框架性能进行综合评价。结果表明, 与单一平台的文生图基线方法(仅使用Midjourney)相比, 完整的3层工作流使FID降低约30%, 专家评定的文化真实性与空间逻辑性提升1.8~1.9分, 同时将人工后期编辑时间减少约57%。分层消融实验显示, 空间表现的提升主要归因于拓扑约束层, 而文化真实性的增强则主要由本体优化层驱动。本研究既建立了一种可复现、与模型无关的文化遗产设计工作流, 也在理论层面搭建了连接当前二维生成范式与未来三维空间认知之间的桥梁。

关键词

中国古典园林; 人工智能; 人工智能生成内容; 设计框架; 设计工作流; 人工智能工具

文章亮点

- 指出人工智能在设计中主要依赖视觉构图, 但缺乏深层空间理解能力
- 指出人工智能对中国古典园林具有基础认知, 但在空间布局与园林要素组织方面仍存在不足
- 提出了一种融合空间布局与园林要素的AIGC辅助设计工作流
- 构建了一种逻辑驱动的多平台协同AIGC设计框架

基金项目

- 国家自然科学基金面上项目(编号: 32371940)
- 教育部人文社会科学研究一般项目(编号: 23YJCZH119)