Generative AI-Based Night Scene Design of Historic Districts: Value Transition from Basic Lighting to Cultural Narrative

Minfei Ran[1], Xiaoyu Lin[1], Mengxiao Tian[1], Jiayi Cong[1], Shuhan Chen[1], Tongguo Wang[1], Yufei Jiang[1]*

[1] Harbin Institute of Technology Shenzhen, Shenzhen, China - universe_7y@naver.com, linxiaoyu@hit.edu.cn, tianmengxiao@hit.edu.cn, congjiayi98@gmail.com, csh0828_2020@qq.com, aryagogo@outlook.com, 2916905853@qq.com
*corresponding author: linxiaoyu@hit.edu.cn

Keywords: Generative Artificial Intelligence, Prompt Enginering, Cultural Heritage Preservation, Nightscape Design, Semantic Prompting, Visual Narratives.

Abstract

As a cultural medium that facilitates the renewal and adaptive reuse of historic neighborhoods, nightscape design is undergoing an evolution from basic illumination to the conveyance of rich cultural narratives. The advent of generative artificial intelligence (AI) in the domain of cultural heritage design has precipitated a paradigm shift, wherein instantaneous engineering has emerged as a pivotal conduit between human intent and AI outcomes. This development has culminated in the enhancement of intelligent visual representations of historic buildings. Generative AI techniques present novel opportunities for conventional lighting design in the context of historic interception. The primary challenge lies in generating nocturnal images that exhibit both aesthetic and cultural depth. This paper explores the potential of generative AI to enhance the efficiency of nightscape design, with a focus on achieving a balance between heritage preservation, contemporary visual language, and cultural significance. The present study proposes a "cultural cue classification system" for transforming daytime imagery into nighttime scenes, based on immediate strategies for cultural heritage. A series of iterative artificial intelligence experiments were conducted using tools such as ChatGPT and Stable Diffusion to evaluate the cultural authenticity, ambience, and detail fidelity of the results, based on the cases of —Nantou Ancient Town and Gankeng Hakka Town in Shenzhen. The findings indicate that culturally embedded cues have a substantial impact on the realism, architectural accuracy, and narrative power of AI-generated images. This accelerated strategy establishes a novel methodological framework for culturally-rich, AI-assisted nighttime narratives.

1. Introduction

Against the backdrop of coordinated development between cultural heritage preservation and urban renewal, nightscape design in historical districts has evolved from serving a basic lighting function to emphasizing narrative expression, cultural representation, and immersive visual experiences. As a key medium for conveying the cultural atmosphere of historical neighborhoods, nightscapes have transitioned from single-purpose illumination to rich visual storytelling that integrates stylistic coherence and historical engagement.In recent years, generative artificial intelligence (GAI), represented by tools such as Stable Diffusion and ChatGPT, has demonstrated tremendous potential in visual generation, scene reconstruction, and cultural narration within architectural and landscape contexts. Nightscape design not only constitutes a core component of urban ambiance but also serves as a critical vehicle for expressing cultural identity.(Liu, Z., & Zhao, L., 2021)Today, with the aid of GAI tools such as Doubao and ChatGPT, we are able to generate richly styled, culturally embedded nightscape images driven by text prompts in a short amount of time—representing a disruptive transformation in traditional design methodologies.(Wang, H., & Wang, Z, 2023)
However, the quality of generated outcomes heavily depends on the structure and semantics of the prompts. (Gero, J., & Yu, R., 2022))Particularly in culturally sensitive contexts such as heritage sites, how to use precise prompts to generate images that align with architectural settings, cultural temperament, and night-time atmospheres remains a critical challenge. This study aims to explore the practical application of AI-based image generation in heritage nightscape design, focusing on how prompt structures and stylistic language affect the generated output. It emphasizes a tripartite framework—"style positioning–light-shadow semantics–cultural symbols"—to construct effective prompt combinations and demonstrates the full process from initial generation to iterative optimization and final rendering with comparative results.(Lu, Y., He, J., & Wang, Y., 2023)
Through field-tested nightscape generation cases across multiple heritage architectural styles, this research uses image-generation AI as the primary tool and conducts experiments and analysis from dimensions such as prompt construction logic, stylistic feedback, and cultural alignment. The objective is to offer a visualized, verifiable, and adjustable AI-assisted method for nightscape design in the field of cultural heritage.

2. Research Background

With the rapid advancement of artificial intelligence technologies, AI-based text-to-image generation is increasingly extending into the arts and creative industries, becoming a growing trend in professional design practices. This research explores how AI-based image synthesis and computer vision technologies can be applied to the automatic optimization of ancient architectural day-to-night visual transformation, thereby providing innovative solutions for the digital documentation, analysis, and management of cultural heritage.
Today, diffusion models are becoming increasingly prevalent in image generation tasks. Natural language-driven applications—from text to image—are booming. In this paradigm, language is not only a means of communication but also a constructive tool. It allows the exploration of the

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-1273-2025

relationship between image generation and traditional architectural environments, opening up new possibilities for the digital recording of built heritage.

In recent years, diffusion models have emerged as a dominant framework in the field of generative image synthesis. By progressively reverse-sampling noisy images, these models enable high-quality and high-fidelity reconstructions, widely applied in artistic creation, urban modeling, and architectural visualization. Comparative studies have shown that, relative to earlier GANs, diffusion models offer clear advantages in structural coherence and preservation of fine details.(Ho, J., Jain, A., & Abbeel, P., 2020)At the same time, natural language-driven image generation has witnessed explosive growth. Complex scene images can now be created through simple textual prompts, significantly enhancing accessibility and creative flexibility. Studies on CLIP-based image generation have also demonstrated that the quality of image generation from text depends not only on model architecture but also on deep semantic embedding and multimodal alignment mechanisms.

With the successive emergence of models such as DALL·E, Stable Diffusion, and Imagen, natural language is increasingly becoming the critical interface connecting user intent with visual output, pushing AI visual design into a new stage of “semantic control.”

3. Research Methodology

This study follows a technical framework consisting of “AI image generation – cultural heritage feature extraction – prompt optimization – outcome validation,” adopting a closed-loop research process combining comparative experiments with iterative optimization. The framework is structured around four stages: data collection, feature analysis, multi-round iteration, and effectiveness evaluation.

The process begins with preliminary photo selection, followed by the design of initial prompts for three nightscape styles: “Elegant Lantern Style,” “Lively Market Style,” and “Zen Landscape Style.” These prompts are used to generate night scenes from daytime photographs. Based on the results of the first generation, the prompts are optimized in terms of illuminance, color, layout, and atmosphere, and a second round of testing is conducted. The core image-generation tool used is Doubao AI, which generates images from text-based prompts. Basic preprocessing, such as backlight dimming and desaturation, is applied to the original daytime images to ensure consistent input quality.

In the third round of experimentation, the refined prompt sets are used again to generate night scenes, serving to verify prompt quality and optimize the output. This cycle is repeated until stable and desirable nightscape images are obtained.(Figure 1)

4. Research Process

4.1 Data Collection and Preprocessing: Multimodal Dataset Construction Based on Field Research

4.1.1 Photographic Plan Design: The research team selected two representative heritage sites in Shenzhen—Nantou Ancient Town and Gankeng Hakka Town—as case study locations. The selected buildings represent a variety of types (ancestral halls, residences, shops), structural characteristics (wooden structures, brick-stone constructions, glazed tiles), and decorative elements (wood carvings, grey sculptures, colored paintings).

Shooting conditions were strictly controlled. Daytime photos were taken under clear, cloudless skies between 9:00 AM and 4:00 PM, using fixed camera settings and aligning the lens with the building’s central axis to ensure consistent angle and height. Night photography was conducted between 6:30 PM and 10:00 PM using the same fixed camera, while also recording ambient lighting (e.g., streetlamps, landscape lighting) and architectural lighting sources (e.g., eave lanterns, interior light seepage).

To ensure sufficient data volume, about 200 pairs of daytime and nighttime images were taken at each site, totaling about 400 images.

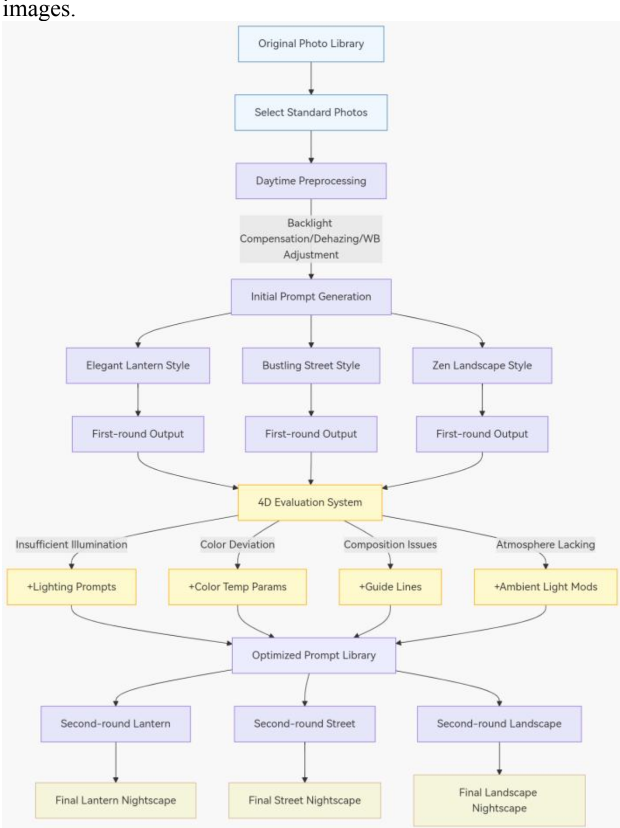

Figure 1:Experimental flowchart

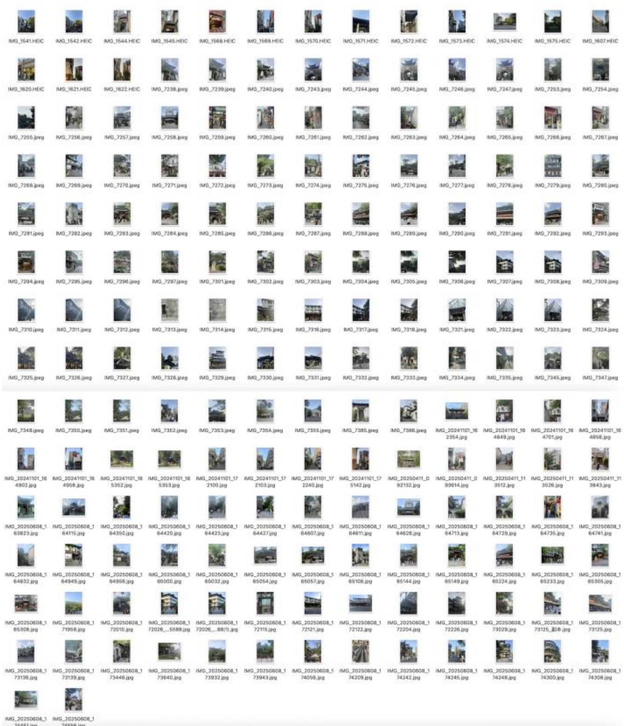
Figure 2:184 daytime images

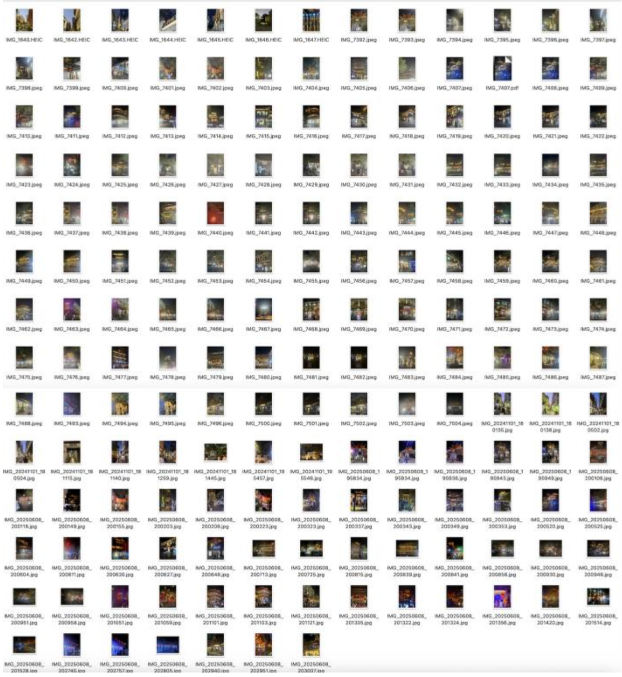

Figure 3:176 nighttime images

4.1.2 Data Preprocessing： First, images were aligned and standardized. For each building, corresponding daytime and nighttime photos were matched based on visual features to form usable image sets. The images were then categorized based on lighting characteristics—for example, “diffused warm yellow from street lamps” or “cool-toned reflections on tiled roofs”—and paired with appropriate style prompts accordingly.

Based on the field research of nightscape photographs, the night view designs of historic districts can be categorized into three distinct styles: Elegant Lantern Style, Lively Market Style, and Zen Landscape Style.

The Elegant Lantern Style is characterized by a warm yellow color palette, featuring hanging lanterns and contour lighting that outlines traditional architecture, evoking a sense of classical elegance. The Lively Market Style employs vivid colors such as warm red and bright yellow, utilizing dense point light sources and linear lighting along signage to create a lively and bustling commercial atmosphere. In contrast, the Zen Landscape Style makes use of cool-toned and low-saturation lighting, complemented by sparsely placed ground lights, fiber optics tracing rooflines, and mist effects projecting ink-wash landscapes, thereby establishing a tranquil and ethereal Zen-like ambiance.

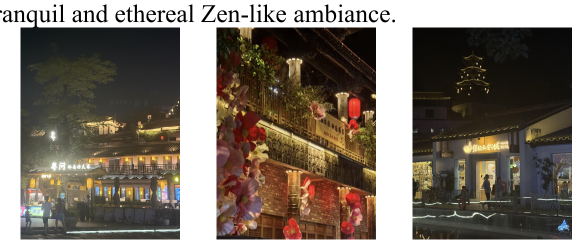

Figure 4: Real shot typical
(From left to right: Style 1-Elegant Lantern Style, Style 2-Lively Market Style, Style 3-Zen Landscape Style)

4.2 Prompt Mining and Optimization

4.2.1 First-Round Experiments: Exploratory Prompt Generation： In this study, the prompts used in the experiments are categorized into two types: “simplified” and “detailed”, which are applied flexibly according to the complexity of the target image and the characteristics of AI responses. The simplified prompts should not be regarded as a “weakened” version; rather, they are a strategic choice that emphasizes “focused intent and reduced detail” at appropriate stages of the generation process. In contrast, the detailed prompts emphasize “fine control and component-specific constraints,” making them more suitable for generation tasks that demand higher output quality. These two types of prompts complement each other, serving different roles at various stages of image generation, and together enhance the accuracy and cultural relevance of AI-generated nightscape imagery.

Simplified prompts are suitable for initial composition and stylistic exploration, focusing on the overall atmosphere, light-color relationships, and spatial perception. They allow AI to quickly produce nightscapes imbued with cultural imagery. On the other hand, detailed prompts are appropriate for stages requiring high-precision control, concentrating on architectural components, light source types, illumination parameters, and other specific elements. These help improve the structural fidelity and expressive quality of the generated images.

Across different AI models, simplified prompts facilitate rapid iteration and are well-suited to platforms with constraints on text length. Detailed prompts enhance the AI’s interpretive accuracy and are more applicable in multi-round interactions or tasks that demand fine-grained control. The combined use of both prompt types effectively supports the multi-stage progression from initial stylistic definition to high-quality image output in AI-assisted nightscape generation.

Figure 4:Selected Photo 1 (Left:Daytime, Right:Nighttime)

Style 1: Elegant Lantern Style

Simplified Prompt: Generate a night scene of an ancient town arcade with warm yellow lanterns symmetrically hung along balconies and eaves; hidden LED strip lights outline dougong structures; ink-style landscape patterns are projected on the walls; warm 3000K lighting blends with greenery to evoke a refined Qinhuai aesthetic.Effect: Symmetrical lantern placement was achieved, but dougong detailing was blurry, ink projections did not scale well with wall surfaces, and wood texture lighting lacked layering.

Detailed Prompt: Emphasize multi-level wooden arcades with carved lattice windows and dougong as focal elements. Install warm yellow lantern-like point lights at eaves and balcony railings; use highlight lighting to render wood carvings; place 3000K LED strip lighting along dougong contours; project 4K ink landscapes on lower walls (projection ratio 1:1.2); retain low-illuminance shadows at the base; maintain central axis symmetry; allow warm lighting to filter through foliage and form halos to convey elegance and tranquility.Effect: Improved clarity of dougong contours, but dynamic blurring of ink projections and natural halo interaction with greenery remained weak.

Figure 5: Style 1-Elegant Lantern Style
(Left:Simplified Prompt, Right:Detailed Prompt)

Style 2: Lively Market Style

Simplified Prompt: Brighten snack and craft stall zones with spotlights; overlay orange transitional lighting on arcades; compress shadow zones to retain a hazy ambiance. Use warm reds (like folk lanterns) and bright yellow signage. Simulate flickering effects with stall lights. Employ dense point lights (stall/ornamental), line lights (neon strips for signage), and area lights (lightboxes for cultural posters) in a casual layout to evoke a vibrant night market.Effect: Initial market atmosphere achieved, but shadow layering was shallow, neon arrangement messy, and flicker effects weak.

Detailed Prompt: Generate a lively market style night scene with bright lighting for street vendors, orange lights on arcades, compressed shadows, warm red and bright yellow tones, subtle flickering, dense light source layout, and freeform positioning for signage and posters to evoke street-level vibrancy.Effect: Slight improvement in atmospheric haze, but poster projection resolution was low and overall lighting contrast weak.

Figure 6: Style 2-Lively Market Style
(Left:Simplified Prompt, Right:Detailed Prompt)

Style 3: Zen Landscape Style

Simplified Prompt: Generate a night scene with 1800K moonlight-like lighting on rooftops, eaves, and old trees. Apply soft blue-gray light to façades; retain 10–15% shadow zones for contemplative voids. Use indigo (#1A237E) and silver-white as low-saturation primary colors. Add dynamic light sway (1–2Hz). Point lights are warm yellow ground lamps spaced 5–8m; fiber lines trace rooftops; fog projections cast 4K ink landscapes freely.Effect: Cold atmosphere achieved, but lamp spacing uneven, fog projection lacked blur, and shadow distribution imbalanced.

Detailed Prompt: Use 1800K lighting to softly highlight rooftops and trees; apply blue-gray light for smooth facade transitions; compress shadows at the base and shaded alleys. Indigo and silver-white simulate moonlight and water reflections. Lights should sway subtly (1–2Hz). Ground lamps (2200K, 5–8m apart) minimize glare; 2mm fiber lines outline rooftops; fog projectors cast 4K ink (1:1.5 ratio, 10% blur) along the building’s curvature. Add sounds of water and rustling bamboo for immersive serenity.Effect: Cold/warm contrast was visible; however, fiber lighting was intermittent and projections lacked spatial coherence.

Figure 6: Style 3-Zen Landscape Style
(Left:Simplified Prompt, Right:Detailed Prompt)

These first-round prompts, while divided into simplified and detailed forms, all exhibited room for improvement. Therefore, second-round experiments focused on prompt refinement.

Differences Between Simplified and Detailed Prompts:The core logic of simplified prompts lies in “focusing on central themes while de-emphasizing details.” By combining key descriptive phrases, simplified prompts quickly define the overall scene style, basic lighting-color relationships, and spatial atmosphere. For example, when generating a nightscape in the Elegant Lantern Style, a simplified prompt may only specify essential elements such as “warm yellow lanterns symmetrically suspended” and “ink-wash projections,” while downplaying specific technical parameters like illuminance levels or distances between light sources. This approach enables the rapid creation of stylistic prototypes.In contrast, detailed prompts adopt a strategy of “component-based constraints and precise control.” These prompts use structured semantic input—such as specific color temperature values, projection ratios, or dynamic frequency settings—to guide the AI model with technical precision. For instance, a detailed prompt for the Zen Landscape Style might include instructions such as “1800K cool lighting to illuminate roof ridges,” “fiber optic diameter of 2mm,” and “10% motion blur,” thereby ensuring a high degree of detail fidelity in the generated image.

Differences in Generated Outcomes:Simplified prompts offer faster generation speeds and are well-suited for stylistic exploration and preliminary composition. They can quickly convey cultural imagery; however, they tend to produce less defined details (e.g., blurred dougong ornamentation) and higher randomness in lighting distribution (e.g., inconsistent spacing between ground lights), requiring subsequent iterations for refinement.Detailed prompts, on the other hand, result in more accurate reconstruction of architectural structures (e.g., highlights on carved wooden windows) and richer lighting hierarchies (e.g., controlled proportions in shadowed areas). Cultural symbols (e.g., intangible heritage posters) are rendered with greater clarity. However, these prompts require longer processing times and place higher demands on the AI model’s comprehension capabilities.In the early design phase, simplified prompts facilitate rapid trial and error to establish the overall stylistic direction (e.g., deciding between Zen Landscape or Lively Market styles). In the optimization phase, detailed prompts support multi-round refinement—such as adjusting the “continuity of fiber optics along the roofline” or the “motion blur of mist projections”—to gradually approach the desired visual outcome.

In summary, simplified prompts yield results with greater “symbolic imagery,” while detailed prompts align more closely with the “precision and narrative depth” expected in professional design.

4.2.2 Second-Round Experiments and Optimized Prompt Effects:

Style 1-Elegant Lantern Style (Optimized):

Optimized Prompt: Generate an arcade night scene with clearly defined warm yellow lanterns symmetrically hung; highlight wood lattice windows with precise lighting; embed 3000K LED strips along dougong structures; project 4K ink landscape (1:1.2 ratio) on walls; retain shadows at the base; warm light should filter through greenery, forming natural halos; ensure symmetry to reinforce a refined Qinhuai style.Effect: Notable improvements in window and dougong detailing, better ink projection alignment, and more natural lighting halos.

Style 2-Lively Market Style (Optimized):

Optimized Prompt: Focus on vibrant marketplaces; use spotlights to highlight food and craft stalls; apply warm orange light for smooth transitions; compress shadows for a hazy effect. Dominant colors: warm red (folk lanterns) and bright yellow (signs). Stall lights flicker subtly. Point lights densely scattered per market flow; neon lines outline signs; lightboxes display intangible heritage posters. Layout remains spontaneous and street-aligned.Effect: Better alignment with market flow; however, warm red was still too dim and flicker effects needed enhancement.

Style 3-Zen Landscape Style (Optimized):

Optimized Prompt: 1800K lighting highlights rooftops, eaves, and trees; blue-gray soft light transitions façades; retain 10–15% base shadow. Main colors: indigo and silver-white; light sways at 1–2Hz. Point lights (warm yellow, 5–8m spacing), 2mm fiber lines for contours, fog-based 4K ink projections (1:1.5, 10% blur) adjust to architectural curves. Ensure harmony between light, architecture, and nature.Effect: Even lamp spacing, smooth fiber lines, natural curvature of projections, and improved cold/warm contrast emphasized serenity.

These optimized prompts yielded significant improvements:

Cultural Authenticity: Stronger expression of architectural period, features, and symbolism enhanced the overall cultural fidelity of generated images.Lighting Design: Prompt details like illuminance, color temperature, dynamic lighting, and spatial division enhanced depth, realism, and artistic presentation.Detail Accuracy: Focused prompt instructions addressed previous issues such as blur, inconsistent scale, and lighting incoherence.Atmosphere and Semantics: Carefully embedded cultural symbols and poetic descriptors enhanced emotional resonance and narrative quality.

A general template emerged for effective prompt structuring: Style Positioning + Lighting & Contrast (Color Temp + Zoning) + Color Strategy (Primary Hues + Dynamic Effects) + Layout Form (Light Types + Distribution Logic) + Atmosphere Design (Cultural Symbolism + Intentional Keywords)

This structured model provides professional-level prompts for cultural heritage nightscape generation, though micro-detailing and dynamic effects still require post-processing or enhancement.

4.2.3 Third-Round Experiment: Validation of Optimized Prompts in Nightscape Generation: Building upon the findings from the first two rounds of experimentation, the third round aimed to further validate the stability and cultural expressiveness of AI-generated heritage nightscape images using optimized prompt strategies.

Addressing the two core challenges identified earlier—(1) lack of cultural contextual sensitivity and visual style specificity in prompts, and (2) the absence of a heritage-oriented prompt structure—this round focused on evaluating prompts in terms of:Cultural Authenticity: Whether the generated image accurately reflects architectural and cultural features.Visual Consistency: The coherence and logical transition from daytime to nighttime imagery.Technical Controllability: The degree of precision with which prompts influence visual outcomes. Stylistic Versatility: Adaptability of prompts to generate varied nightscape styles.

Sample images and results from this round demonstrated:

Style 1-Elegant Lantern Style: Improved lantern symmetry; dougong details still required enhancement.

Style 2-Lively Market Style: Strong market ambiance; further refinement needed for lighting layout.

Style 3-Zen Landscape Style: Tranquil mood achieved; micro-detailing still to be improved.

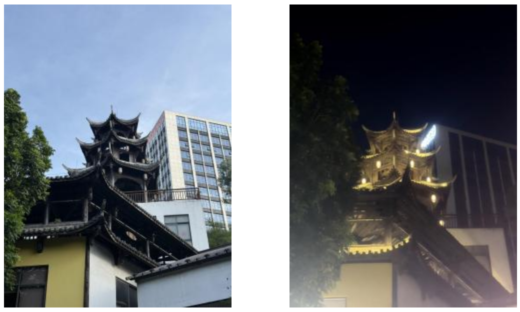

Figure 7: Selected Photo 2 (Left:Daytime, Right:Nighttime)

Figure 8: Style 1-Elegant Lantern Style(Lantern symmetry is well-achieved, but dougong detailing still requires enhancement.)

Fighre 9: Style 2-Lively Market Style(Atmosphere is vibrant, yet light source distribution needs further refinement.)

Fighre 10: Style 3-Zen Landscape Style (Tranquility is well conveyed, but scene details demand further improvement.)

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-1273-2025 | © Author(s) 2025. CC BY 4.0 License.

Summary of Round 3 Outcomes:

Style Type	Round	Core Prompt Content	Generation Outcome	Optimization Direction
Elegant Lantern Style	Round 1	Simplified Version: Warm yellow lanterns symmetrically suspended; dougong (bracket sets) outlined with light strips; ink-wash landscape projected on walls; 3000K warm light blended with greenery. Detailed Version: Emphasizes wooden arcades and carved window details; point light sources at eaves lanterns and linear LED light along dougong; 4K ink-wash projection (throw ratio 1:1.2); symmetrical central-axis layout; warm light permeates greenery, creating a glowing halo.	Lanterns are symmetrically placed, but dougong details are vague; projection scale mismatched with walls; weak wood texture layering in lighting.	Enhance dougong highlight details, adjust projection ratio, improve natural blending of halo effects between warm light and greenery.
	Round 2	Based on Round 1, specifies light strips outlining the dougong structure, adjusts projection ratio to 1:1.2, strictly follows central-axis symmetry, and emphasizes carved window details and light-shadow layering.	Improved details in dougong and window carvings, better projection alignment, yet lack of dynamic blur effect in ink-wash projection.	Add dynamic blur parameters to projection; enhance interaction between light sources and greenery.
	Round 3	Utilizes prompt content from Round 2.	Good lantern symmetry and natural lighting layers, but dougong carving details still need enhancement.	Focus on microstructures (e.g., dougong mortise and tenon joints); increase realism in architectural detailing.
Lively Market Style	Round 1	Simplified Version: Highlighted snack street with strong key lighting; warm orange tones transition across arcades; compressed shadow zones evoke hazy ambiance; main palette of warm red with bright yellow signage; flickering stall lights; free-form use of point/linear/area light sources. Detailed Version: Emphasizes "night market with a thousand lights"; highlights density of light sources matching public movement lines; heritage poster lightboxes using area lighting.	Initial festive ambiance achieved, but shadow layering is flat; neon and signage layout is cluttered; stall lights lack dynamic flicker effects.	Refine light source distribution logic; enhance clarity of heritage elements; improve realism of dynamic lighting effects.
	Round 2	Focuses on "urban movement line" light arrangement: point lights (stall lamps) dispersed in high-traffic zones; linear lights (neon) outline signage; area lights (lightboxes) display heritage posters; warm red tones mimic traditional lantern colors.	Light source distribution better matches scene logic; warm red tones appear too dim; stall light flickering still not distinct.	Adjust color temperature parameters; refine dynamic lighting effects such as flicker frequency and brightness variations.
	Round 3	Utilizes prompt content from Round 2.	Vibrant ambiance achieved, light arrangement follows commercial flow, but resolution of heritage poster projection remains low.	Increase resolution of area light projections; add environmental audio-visual interaction parameters.
Zen Landscape Style	Round 1	Simplified Version: 1800K cool lighting highlights rooftops and eaves; bluish-gray soft light transitions across facades; 10–15% shadow left at the base; primary tones of indigo and silver-white; ground lamps spaced 5–8m apart; fog projections of ink-wash landscapes. Detailed Version: Cool lighting simulates moonlight; ground lamps with 2200K	Atmosphere of cool light achieved, but ground lamp spacing uneven; fog projection lacks dynamic blur; shadow-to-light balance inconsistent.	Standardize spacing of ground lamps; introduce dynamic projection parameters; optimize shadow-white space ratio.

		color temperature; 2mm fiber optic outlines rooftops; fog projection with 10% dynamic blur; integrates poetic imagery of "a secluded mountain after new rain."		
	Round 2	Precisely controls 1800K cool light temperature and lamp spacing; uses fiber optics to outline rooftops continuously; ink-wash projection with 1:1.5 throw ratio and 10% dynamic blur; emphasizes "coexistence of light-shadow with heritage architecture and greenery."	Even lamp distribution; projection conforms to curved wall surfaces; fiber outlines still appear disconnected; contrast between cool and warm lighting is weak.	Increase color temperature contrast; enhance continuity of fiber optic outlines.
	Round 3	Utilizes prompt content from Round 2.	Strong sense of tranquility; integration of light and natural elements achieved; micro-details (e.g., texture of ancient trees) still need refinement.	Add parameters for natural texture shading to enhance immersive experience.

Ultimately, a generalized prompt structure was established:Style Positioning + Lighting and Contrast (Color Temperature + Light-Zone Division) + Color Strategy (Primary Color Values + Dynamic Effects) + Spatial Layout (Light Type + Distribution Logic) + Atmospheric Design (Cultural Symbols + Intent Keywords)

This structure was found to be broadly applicable for professional-grade nightscape generation in cultural heritage contexts. However, micro-level realism and complex dynamic effects still benefit from post-processing techniques or hybrid workflows.

5. Conclusion and Outlook

This study systematically explored the value and methodology of prompt engineering for nightscape image generation in cultural heritage settings. A structured prompt model was proposed and experimentally validated across multiple architectural heritage styles. Results confirmed that prompt style directly impacts the generated image's cultural expressiveness, color accuracy, and immersive atmosphere.

The study proposed an optimized prompt framework centered on structured semantic control:Style Positioning + Light-Zone Semantics (Color Temp + Distribution) + Color Representation (Hue Values + Dynamics) + Spatial Form (Light Type + Layout) + Cultural Atmosphere (Symbols + Poetic Language)

Empirical case studies demonstrated that incorporating style indicators (e.g., historical era, regional craftsmanship), lighting semantics (e.g., color temperature, focal illumination), and cultural imagery (e.g., festival lanterns, folk patterns) significantly improved the fidelity and narrative depth of AI-generated visuals.

This paper also documented the iterative generation process—from vague initial outputs to increasingly refined results via semantic prompt calibration. Comparative analysis confirmed that structured prompts enhanced visual stability and stylistic coherence, particularly in nightscape generation.

However, limitations remain. First, the dataset size may constrain generalizability across architectural types and lighting conditions. Second, prompt construction was largely manual and lacked adaptive intelligence. Third, although visually expressive, the generated outputs lacked systematic evaluation in cultural relevance, semantic alignment, and perceptual comfort.

Future research should prioritize:Large-Scale Heritage Image Corpora: Constructing a dataset with semantic tags and style annotations to support automated prompt generation. AI-Assisted Prompt Recommendation: Developing intelligent engines to enhance prompt efficiency and personalization. (Lee, K., Kim, S., & Park, J. 2023)Expanded Application Scenarios: Extending this methodology to digital museums, virtual tourism, cultural performances, and other immersive settings—advancing the fusion of culture and technology through multimodal heritage expression and cross-disciplinary dissemination.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No.52208018,No.52408011); the Guangdong Philosophy and Social Science Planning Project (2023 Discipline Co-construction Project) (No. GD23XLN29)；by Annual project of Harbin Institute of Technology (Shenzhen)（No. FYPP2024F0136）

References

[Anonymous]. (2024). Development of a Method for Commercial Style Transfer of Historical Architectural Facades Based on Stable Diffusion Models. Journal of Imaging, 10(7), 165.

Gero, J., & Yu, R. (2022). Artificial intelligence in architecture: Generative models and prompt engineering. Design Studies, 76, 101072.

Guo, F., Liu, J., Xie, Q., Chang, H. (2023). Improved Nighttime Traffic Detection Using Day-To-Night Image Transfer. Transportation Research Record, 2677(5), 1–12.

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. arXiv preprint arXiv:2006.11239.

Lee, K., Kim, S., & Park, J. (2023). Semantic prompting for AI-generated cultural scenes. Journal of Cultural Heritage, 61, 117–124.

Liu, Z., & Zhao, L. (2021). From illumination to narrative: The evolution of nightscape design in historic districts. Architectural Journal, (5), 44–49.

Lu, Y., He, J., & Wang, Y. (2023). Prompt engineering in cultural heritage AI applications: A structured semantic approach. Heritage Science, 11(1), 35.

Wang, H., & Wang, Z. (2023). Research on generative AI-assisted mechanisms in architectural design. Architectural Journal, (2), 70–76.