



EVALUATING GENERATIVE AI MODELS FOR HISTORIC URBAN HERITAGE VISUALIZATION IN SMART CITIES

A CASE STUDY OF AL-MU'IZZ STREET FOR SMART CITY PRESERVATION

RENADA ABD ALKADER¹, ABDALLA MOUSTAFA¹

¹ KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS, SAUDI ARABIA

KEYWORDS	ABSTRACT
<i>Urban Heritage</i> <i>Generative Artificial Intelligence</i> <i>Heritage Conservation</i> <i>Smart Cities</i> <i>Al-Mu'izz Street</i>	<i>Evaluating generative AI models across five sites, this study investigates historical urban heritage visualization on Cairo's Al-Mu'izz Street to advance smart city preservation frameworks. Findings demonstrate that Gemini 3.1 Pro achieves superior conservation fidelity (84.9% overall; 4.20/5 in retention and historical reliability), whereas ChatGPT 5.6 Sol excels in visual quality and material expression. Cross-model optimization reveals that ChatGPT-derived instructions paired with Gemini execution yield the highest authenticity (89.8%). Ultimately, integrating structured, AI-assisted prompt workflows into digital twin environments provides urban planners with a scalable, historically reliable mechanism for heritage conservation and urban management.</i>

1. Introduction

Historic Cairo, designated as a UNESCO World Heritage Site in 1979, contains one of the highest concentrations of Islamic architectural monuments in the world (Abdelalim, 2025). The primary structural spine of this medieval city is Al-Mu'izz li-Din Allah al-Fatimi Street (historically known as the Qasaba), a one-kilometer pedestrian corridor founded in the 10th century by the Fatimids (Abdel-Hadi et al., 2009; *Moez Street in Cairo*, 2024). Over successive centuries, Fatimid, Ayyubid, Mamluk, and Ottoman rulers constructed monumental religious, educational, and philanthropic complexes that shaped the street's organic, tightly enclosed medieval morphology (Abdel-Hadi et al., 2009; *Moez Street in Cairo*, 2024).

Among these structures, the Sabil-Kuttab typology, such as the Sabil-Kuttab of Khusru Pasha (1535 AD) and the Sabil-Kuttab of Abd al-Rahman Katkhuda (1744 AD), served a dual philanthropic function, supplying free public drinking water on the ground floor while housing a Quranic elementary school above (Abdelalim, 2025). For centuries, these monuments held the historic urban fabric in place (ElKatsha, 2000). However, modern urban pressures, environmental decay, and late 20th-century interventions, such as the bisection of the corridor by Al-Azhar Street in 1931 and top-down pedestrianization in 2008, have altered the social and physical fabric of Al-Mu'izz Street, prompting ongoing debates regarding aesthetic authenticity versus the "counterfeit" commercialization of historic urban spaces (Abdel-Hadi et al., 2009; ElKatsha, 2000).

To address the vulnerabilities of historic urban centers, heritage documentation has evolved from manual inspections toward advanced digital workflows (Abdelalim, 2025). The transition from standard Building Information Modeling (BIM) to Historic Building Information Modeling (HBIM) integrates high-precision 3D Terrestrial Laser Scanning (TLS) point clouds and photogrammetry with semantic historical metadata (Abdelalim, 2025; Mukhamedjanova, 2025). When combined with real-time Internet of Things (IoT) sensors, HBIM forms the foundation of Heritage Digital Twins (HDTs), virtual replicas that reflect a physical monument's real-time condition, support proactive maintenance, and enable risk simulation without risking physical damage (Abdelalim, 2025; Architectours, 2025).

Concurrently, Artificial Intelligence (AI) has expanded across five core conservation phases: structural monitoring, damage detection, microclimate prediction, risk assessment, and environmental control (Tan et al., 2026). Computer vision models automatically identify surface decay and masonry cracks, while Generative AI (GAI) and multimodal large language models enable the automated synthesis of photorealistic 3D architectural scenarios and text-driven visual reconstructions (Mukhamedjanova, 2025; Ran et al., 2025; Tan et al., 2026).

Despite these computational advancements, deploying Generative AI in architectural heritage introduces severe methodological and ethical risks (Ghaith, 2024; Mukhamedjanova, 2025). GAI tools are prone to visual hallucinations, frequently generating speculative structural additions or fake decorative elements that contradict archival documentation (Mukhamedjanova, 2025). Furthermore, because commercial models are predominantly trained on Western datasets, they exhibit algorithmic and regional biases, often overwriting complex non-Western vernacular details, such as Cairene ablaq stone masonry, wrought-iron window grilles, or carved wooden mashrabiya, with generic, homogenized "Middle Eastern" clichés (Abdelalim, 2025).

To overcome these limitations, recent research emphasizes structured prompt engineering, specifically the "Alter/Retain" framework (Sektani et al., 2026). By translating international conservation doctrines into rigid text parameters that isolate permitted modifications ("Alter") from mandatory historical baselines ("Retain"), planners can control visual outputs and evaluate GAI tools across a continuum of preservation scenarios (Sektani et al., 2026). However, an empirical benchmark comparing how leading commercial generative models navigate these strict conservation constraints within dense Cairene urban fabrics remains a critical gap in smart city literature (Mukhamedjanova, 2025).

2. Literature Review

The integration of advanced digital methodologies within cultural heritage conservation marks a significant shift in how historic urban environments are documented, analyzed, and reimaged (Abdelalim, 2025; Ghaith, 2024). As smart city initiatives increasingly seek to incorporate cultural legacy as a core pillar of urban livability and tourism, the role of Artificial Intelligence (AI) has progressed from passive monitoring to active visual generation (Angelidou et al., 2017; Borda & Bowen, 2017). This literature review delves into several key domains: the urban heritage context and morphology of Historic Cairo, the progression of AI in conservation, the mechanics of prompt engineering, the role of generative AI in architectural visualization, and the emerging integration of these workflows within smart city digital twin platforms, ending in a synthesis of the current research gap.

2.1. Urban Heritage Context and the Morphology of Historic Cairo's Al-Mu'izz Street

Historic Cairo, recognized as a UNESCO World Heritage Site in 1979, represents one of the most historically dense and architecturally significant urban sites in the Islamic world (Abdelalim, 2025). The primary structural spine of this medieval gated city is Al-Mu'izz li-Din Allah al-Fatimi Street (historically known as the Qasaba) (Abdel-Hadi et al., 2009; "Al-Mu'izz Street – The Heart of Islamic Cairo," 2026; *Moez Street in Cairo*, 2024). Founded in the 10th century by the Fatimid Caliph al-Mu'izz, the street served as a highly prestigious, ceremonial, and commercial corridor running north-to-south from Bab al-Futuh to Bab Zuwayla (Abdel-Hadi et al., 2009; *Moez Street in Cairo*, 2024). Over successive centuries, each ruling dynasty, including the Ayyubids, Bahri and Burji Mamluks, and the Ottomans, constructed monumental structures that incrementally shaped the street's dense urban morphology (Abdel-Hadi et al., 2009; *Moez Street in Cairo*, 2024; *Streets of Islamic Cairo: A Configuration of Urban Themes and Patterns*, n.d.).

The spatial layout of Al-Mu'izz Street is defined by its medieval urban morphology, characterized by a narrow, winding centerline, spontaneous variations in width and building heights, and sequential visual vistas that frame towering minarets and recessed monumental facades (Abdel-Hadi et al., 2009). This organic urban fabric was held in place for centuries by major religious and philanthropic complexes (ElKatsha, 2000). A key typology within this fabric is the Sabil-Kuttab, a unique Cairene institution combining a ground-floor public water fountain (sabil) with an upper-floor elementary school (kuttab), representing a masterful synthesis of Ottoman and Mamluk architectural traditions (Abdelalim, 2025).

However, modern interventions have introduced severe physical and social transformations (ElKatsha, 2000). The construction of major east-west vehicular axes, such as Al-Azhar Street in 1931, severed the historic north-south pedestrian spine, disrupting the continuous rhythm of monuments (ElKatsha, 2000). Furthermore, major rehabilitation efforts initiated in the late 1970s culminated in the pedestrianization of northern Al-Mu'izz Street in 2008 to create an "open museum" (Abdel-Hadi et al., 2009). This top-down pedestrianization replaced traditional, multi-layered Cairene street life, characterized by close-knit local sociability, crafts, and micro-trades, with a sanitized, highly commercialized tourist zone, sparking intense academic debates regarding aesthetic authenticity versus the "counterfeit" sterilization of historic districts (Abdel-Hadi et al., 2009; "Al-Mu'izz Street – The Heart of Islamic Cairo," 2026; ElKatsha, 2000).

2.2. Artificial Intelligence in Cultural Heritage Conservation

Within the broader domain of heritage conservation, Artificial Intelligence has emerged as a powerful tool to address the limitations of traditional manual inspection (Ghaith, 2024; Tan et al., 2026). Research indicates that AI applications can be systematically classified across five key preservative phases (Tan et al., 2026):

- **Monitoring:** Utilizing IoT sensors and video feeds to track structural health, groundwater levels, and indoor microclimates (Artchitectours, 2025; *Streets of Islamic Cairo: A Configuration of Urban Themes and Patterns*, n.d.).

- Damage Identification and Detection: Applying computer vision to detect stone cracks, masonry decay, and surface deterioration (*Streets of Islamic Cairo: A Configuration of Urban Themes and Patterns*, n.d.; Tan et al., 2026).
- Environmental Prediction: Using machine learning to anticipate changes in temperature and humidity (Tan et al., 2026).
- Deterioration Risk Assessment: Processing multi-dimensional datasets to calculate structural risk indexes (Tan et al., 2026).
- Environmental Control: Driving intelligent, adaptive climate controls in historic structures to prevent material decay (Tan et al., 2026).

While predictive algorithms and neural networks have greatly enhanced diagnostic precision, the introduction of Generative AI (GAI) for visual restoration and reconstruction has introduced major ethical complexities (Ghaith, 2024; Mukhamedjanova, 2025). Standard conservation doctrines, codified in international charters, stress the principles of minimal intervention, reversibility, and the preservation of historic patina (Ghaith, 2024; Mukhamedjanova, 2025). GAI models, however, are highly prone to visual hallucinations, wherein they generate historically inaccurate decorations or speculative structural elements (Mukhamedjanova, 2025). Furthermore, because commercial models are trained on highly generalized, Western-centric architectural datasets, they carry strong stylistic and regional biases, often overwriting complex non-Western vernacular details, such as Cairene ablaq stone courses or wooden mashrabiya, with generic, homogenized Middle Eastern design clichés (Abdelalim, 2025; Mukhamedjanova, 2025).

2.3. Prompt Engineering and Heritage Visualization

To mitigate the speculative nature of GAI, researchers have increasingly turned to prompt engineering as a mechanism of strict semantic control (Ran et al., 2025). In this paradigm, natural language functions not just as an input, but as a precise design tool that guides image synthesis (Ran et al., 2025). A major advancement in this domain is the "Alter/Retain" prompt engineering framework (Sektani et al., 2026). This methodology translates abstract conservation intensities into highly structured visual parameters (Sektani et al., 2026). Prompts are split into two distinct, rigid parameters (Sektani et al., 2026):

- The "Alter" Parameter: Explicitly defines the targeted, permissible physical modifications allowed under specific conservation guidelines (Sektani et al., 2026).
- The "Retain" Parameter: Dictates which original architectural assets, geometries, scales, materials, and contextual details must remain completely untouched to maintain historical integrity (Sektani et al., 2026).

To achieve authentic historical rendering, researchers have proposed structured cultural cue classification systems (Ran et al., 2025). By segmenting prompts into distinct style positionings, light-zone semantics (defining color temperature and distribution), and cultural symbols, GAI models are forced to produce structurally and atmospheric coherent results (Ran et al., 2025). Furthermore, integrating GAI with Low-Rank Adaptation (LoRA) fine-tuning allows models to learn highly specialized, localized typological data, making them context-aware and capable of rendering realistic physical finishes (such as weathered Cairo limestone or dark basalt interlocked paving) rather than producing overly polished, cartoonish visual outputs (Abdel-Hadi et al., 2009; Lee & Park, 2025).

2.4. Generative AI in Architectural Visualization within Smart City Digital Twin System

The rapid evolution of GAI has fundamentally altered the field of architectural visualization (*AI in Architectural Concept Design: Reimagining the Conflict-Affected Areas in Syria*, n.d.; Tan et al., 2026). Traditionally, representing alternative future scenarios for deteriorated structures required labor-intensive, static 3D modeling workflows (Abdelalim, 2025). GAI now enables

rapid, iterative visualization of proposed conservation interventions, allowing designers to easily represent a graded continuum of preservation options (Sektani et al., 2026).

In modern smart city planning, these visual capabilities are increasingly being integrated with Heritage Digital Twins (HDTs) and Historic Building Information Modeling (HBIM) (Abdelalim, 2025; Mukhamedjanova, 2025). Unlike standard BIM, which is optimized for new construction, HBIM embeds historical metadata, material properties, and degradation tracking directly into semantic 3D geometries generated via Terrestrial Laser Scanning (TLS) and photogrammetry (Mukhamedjanova, 2025). An HDT acts as a dynamic virtual replica that evolves in sync with the physical monument through IoT sensors (Abdelalim, 2025; Architectours, 2025).

By aligning GAI-generated visual scenarios with the high-precision geometries of HDTs, urban planners can establish highly immersive platforms for participatory urban planning and community engagement (*AI in Architectural Concept Design: Reimagining the Conflict-Affected Areas in Syria*, n.d.; Lee & Park, 2025). Abstract conservation and adaptive reuse proposals can be translated into easily interpretable, visually realistic alternatives (Abdelalim, 2025; Sektani et al., 2026). This allows local stakeholders and non-experts to intuitively understand spatial transformations during workshops, bridging the gap between professional conservation theory and the lived realities of the local community (*AI in Architectural Concept Design: Reimagining the Conflict-Affected Areas in Syria*, n.d.; Lee & Park, 2025; Sektani et al., 2026).

2.5. Synthesizing the Research Gap

Despite the clear convergence of these fields, a critical research gap remains in the existing literature:

1. Lack of Empirical Multi-Model Benchmarking: While the "Alter/Retain" prompt framework has been tested on singular GAI models, there is an absolute absence of comparative research evaluating how different advanced commercial models perform against each other when subjected to the same strict heritage constraints (Sektani et al., 2026).
2. Lack of Evaluation in Non-Western Urban Fabrics: Existing studies focusing on AI prompt stability are primarily tested on Western or East Asian contexts (Ran et al., 2025; Sektani et al., 2026). There is a lack of rigorous, quantitative visual testing evaluating how commercial AI models handle the complex, multi-layered architectural syncretism of Islamic Cairo, where Fatimid, Mamluk, and Ottoman layers exist on a single narrow street centerline (Abdel-Hadi et al., 2009).
3. Methodological Fragmentation: There remains a clear integration gap between overarching smart city digital twin platforms (which focus on engineering-oriented structural monitoring and static GIS coordinates) and site-based creative GAI visualization tools (Angelidou et al., 2017; Mukhamedjanova, 2025). A formalized, reproducible workflow that explicitly links GAI's prompt-driven visualization capabilities with HBIM frameworks for community co-design has yet to be fully established and empirically tested.

To resolve these gaps, this paper utilizes Al-Mu'izz Street and the Sabil-Kuttab of Abd al-Rahman Katkhuda as a testing ground to systematically benchmark ChatGPT Pro and Gemini Pro, establishing an actionable framework for AI-assisted urban preservation in 21st-century smart cities.

3. Methodology

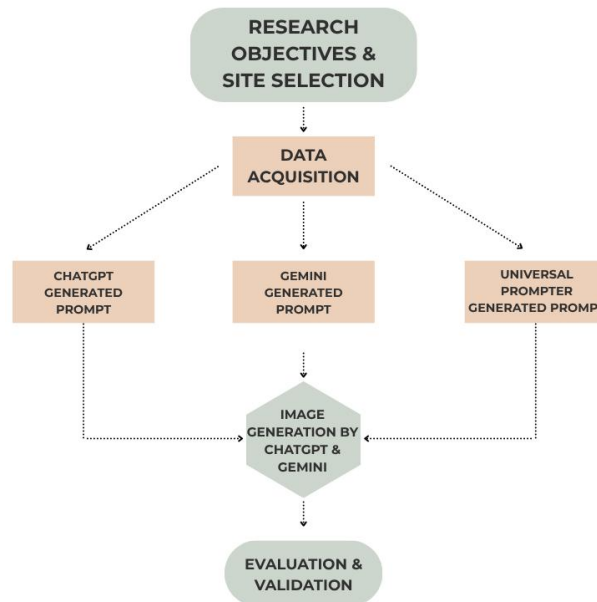
The methodology established follows a systematic framework combining qualitative site analysis with artificial intelligence-driven visual synthesis. The research process was initiated by defining the core aim and objectives of the paper, which established the foundational parameters for the study. Guided by these primary objectives, a specific case study location was selected to ground the theoretical framework in a practical, physical context. To establish a

robust analytical baseline, a comprehensive literature review was conducted alongside the systematic gathering of site-specific images, as thorough documentation is the first step to preserving buildings and structures of great architectural or historical value.

Following the initial data collection, the study transitioned into the computational design phase utilizing natural language-driven applications. The prompt engineering process involved formulating three distinct textual commands using ChatGPT 5.6 Sol, Gemini 3.1 Pro, and Universal Prompter. Because mastering the art of prompting is crucial for producing accurate and visually appealing images, these structured prompts were designed to effectively translate the site's spatial and visual data into specific directives. These three prompts were subsequently deployed across two generative AI platforms, specifically Gemini 3.1 Pro and ChatGPT 5.6 Sol, to execute the image generation phase and synthesize the visual iterations.

The last phase of the research consisted of a structured evaluation stage. The AI-generated visual outputs were critically assessed to determine their accuracy, visual quality, and alignment with the initial objectives. This rigorous evaluation of AI tool designs is crucial in ensuring that the generated concepts meet the predefined criteria. Figure 1 illustrates the methodological flowchart.

Figure 1. Methodology Flowchart.



4. Results

4.1. ChatGPT Visual Results

Figure 2 compares the original Site 1 photograph with three ChatGPT-generated heritage visualization outputs produced using different prompting strategies. Image (a) presents the existing site condition, characterized by the historic mosque façade and dome, the prominent minaret, the surrounding traditional buildings, and the active pedestrian streetscape. Image (b), generated using the ChatGPT-derived prompt, introduces a more curated and visually refined public realm with improved paving, organized pedestrian space, planting, and street furniture, while maintaining the overall heritage character. Image (c), generated using the Gemini-derived prompt, preserves the original architectural composition more closely, particularly the relationship between the dome, minaret, façade, and adjacent building mass, while integrating moderate streetscape improvements. Image (d), generated using the Universal Prompter, creates a more generalized historic urban scene with a lively pedestrian environment and traditional architectural character; however, it introduces greater changes to the original spatial and architectural configuration. Overall, the comparison illustrates how different prompting

strategies influence the balance between visual enhancement, site fidelity, and heritage preservation.

The second and subsequent paragraphs in sections must have 0.5 first line indentation. title.

Figure 2. ChatGPT Image Generation for Site 1: (a) Original Site Image; (b) ChatGPT Generated Photo by ChatGPT Prompt; (c) ChatGPT Generated Photo by Gemini Prompt; (d) ChatGPT Generated Photo by Universal Prompter.



Figure 3 compares the original Site 2 photograph with three ChatGPT-generated visualization outputs produced using different prompting strategies. Image (a) shows the existing site condition, featuring the large historic dome, the stone façade with pointed recessed panels, the projecting wooden mashrabiya-like structure, and the surrounding pedestrian activity. Image (b), generated using the ChatGPT-derived prompt, presents a refined heritage setting with upgraded paving, organized circulation, seating, lighting, planting, and signage, while preserving the main architectural identity of the original structure. Image (c), generated using the Gemini-derived prompt, maintains the original building composition more closely, particularly the dome, façade proportions, and adjacent wooden structure, while introducing a cleaner and more visually integrated public realm. Image (d), generated using the Universal Prompter, transforms the site into a more elongated and monumental heritage streetscape, with stronger urban reinterpretation and additional background architectural elements, resulting in a scene that is visually convincing but less faithful to the original site configuration. Overall, the comparison highlights how different prompting strategies affect the balance between architectural preservation, visual enhancement, and contextual reinterpretation.

Figure 3. ChatGPT Image Generation for Site 2: (a) Original Site Image; (b) ChatGPT Generated Photo by ChatGPT Prompt; (c) ChatGPT Generated Photo by Gemini Prompt; (d) ChatGPT Generated Photo by Universal Prompter.



Figure 4 illustrates the original Site 3 photograph with three ChatGPT-generated heritage visualization outputs produced using different prompting strategies. Image (a) shows the existing site condition, characterized by a narrow historic street, active commercial edges, dense pedestrian movement, and the prominent heritage gateway/building terminating the view. Image (b), generated using the ChatGPT-derived prompt, presents a more organized and visually refined streetscape with improved paving, signage, lighting, and a clearer pedestrian environment, while preserving the central monument and the overall linear composition. Image (c), generated using the Gemini-derived prompt, introduces a more formalized and spatially widened interpretation of the site, with cleaner architectural edges and a curated public realm, while still maintaining the main gateway and historic character. Image (d), generated using the Universal Prompter, creates a lively and visually rich heritage street scene with enhanced pedestrian activity and commercial vitality, but it involves greater reinterpretation of the surrounding urban fabric and skyline. Overall, the comparison demonstrates how different prompting approaches influence the balance between site fidelity, streetscape enhancement, and the preservation of the original heritage character.

Figure 4. ChatGPT Image Generation for Site 3: (a) Original Site Image; (b) ChatGPT Generated Photo by ChatGPT Prompt; (c) ChatGPT Generated Photo by Gemini Prompt; (d) ChatGPT Generated Photo by Universal Prompter.



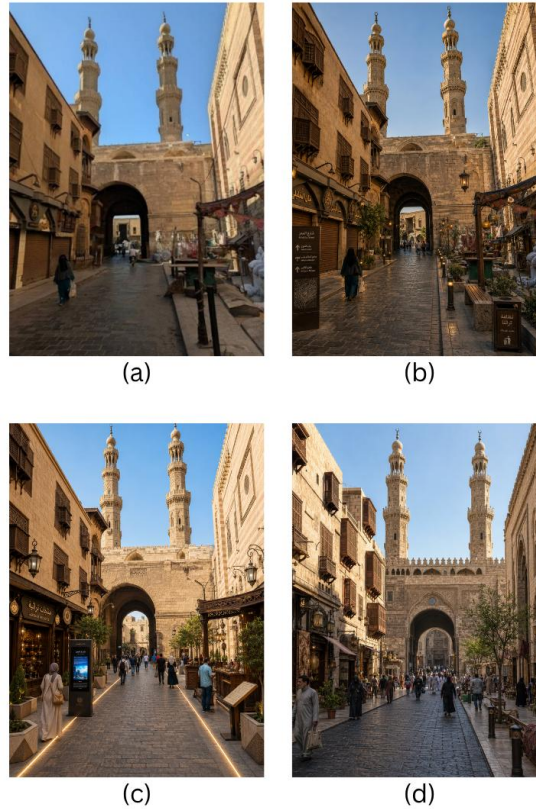
As shown Figure 5, Image (a) shows the existing site condition, featuring the prominent central minaret, the large dome to the left, the secondary minaret in the background, and the surrounding historic street edge. Image (b), generated using the ChatGPT-derived prompt, presents a visually enhanced heritage setting with improved paving, signage, lighting, and pedestrian organization, while retaining the dominant architectural features of the original mosque complex. Image (c), generated using the Gemini-derived prompt, preserves the original composition more closely, particularly the spatial relationship between the minaret, dome, and surrounding buildings, while introducing a cleaner and more structured public realm. Image (d), generated using the Universal Prompter, creates a lively and aesthetically appealing heritage streetscape with added greenery and pedestrian activity, but it involves a somewhat greater reinterpretation of the surrounding built environment.

Figure 5. ChatGPT Image Generation for Site 4: (a) Original Site Image; (b) ChatGPT Generated Photo by ChatGPT Prompt; (c) ChatGPT Generated Photo by Gemini Prompt; (d) ChatGPT Generated Photo by Universal Prompter.



Figure 6 evaluates three ChatGPT-generated visualization approaches against the original photograph of Site 5. (a) The ChatGPT-prompted output (b) retains the fundamental architectural character while introducing a structured, pedestrian-friendly upgrade with improved greenery and paving. Similarly, the Gemini-prompted output (c) maintains the original arch and minarets while delivering an orderly and contemporary public realm design. In contrast, the Universal Prompter output (d) produces a lively, dramatic street scene with expanded pedestrian movement, yet it diverges significantly from the site's authentic configuration. Together, these results demonstrate how different prompting strategies balance historical preservation, public space enhancement, and creative interpretation.

Figure 6. ChatGPT Image Generation for Site 5: (a) Original Site Image; (b) ChatGPT Generated Photo by ChatGPT Prompt; (c) ChatGPT Generated Photo by Gemini Prompt; (d) ChatGPT Generated Photo by Universal Prompter.



4.2. Gemini Visual Result

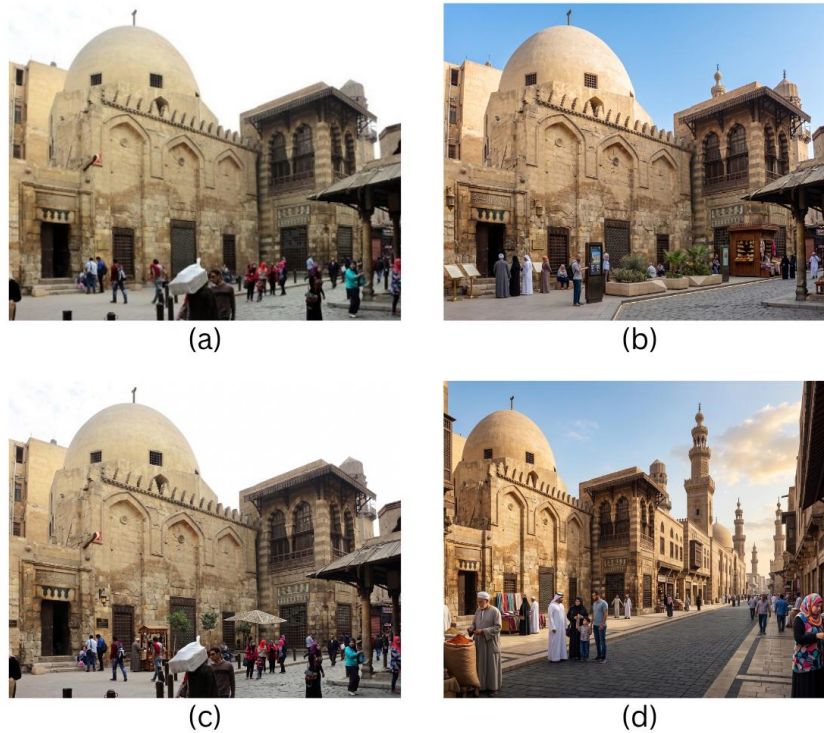
Figure 7 compares the original Site 1 photograph with three Gemini-generated outputs. Image (b), produced using the Gemini prompt, enhances the streetscape while preserving the main dome, minaret, and surrounding heritage fabric. Image (c), generated using the ChatGPT prompt, maintains the original architectural composition more closely with relatively limited intervention. Image (d), generated using the Universal Prompter, also retains the major heritage elements but introduces a more generalized and simplified streetscape treatment. Overall, the three outputs show different balances between visual enhancement and fidelity to the original site.

Figure 7. Gemini Image Generation for Site 1: (a) Original Site Image; (b) Gemini Generated Photo by Gemini Prompt; (c) Gemini Generated Photo by ChatGPT Prompt; (d) Gemini Generated Photo by Universal Prompter.



Figure 8 compares the original Site 2 photograph with three Gemini-generated outputs. Image (b), produced using the Gemini prompt, preserves the main dome, façade, and projecting wooden structure while introducing a more organized and refined public realm. Image (c), generated using the ChatGPT prompt, remains closest to the original site condition with minimal intervention. Image (d), generated using the Universal Prompter, creates a more extensive heritage streetscape with additional architectural elements and a stronger reinterpretation of the original spatial context. Overall, the outputs differ mainly in the degree of preservation versus transformation.

Figure 8. Gemini Image Generation for Site 2: (a) Original Site Image; (b) Gemini Generated Photo by Gemini Prompt; (c) Gemini Generated Photo by ChatGPT Prompt; (d) Gemini Generated Photo by Universal Prompter.



As shown in Figure 9, the original Site 3 photograph with three Gemini-generated outputs. Image (b), generated using the Gemini prompt, presents a more formalized and organized heritage streetscape while preserving the central gateway. Image (c), generated using the ChatGPT prompt, remains closest to the original market character, street activity, and spatial composition. Image (d), produced using the Universal Prompter, maintains a lively pedestrian environment but introduces greater reinterpretation of the surrounding buildings and streetscape. Overall, the main difference lies in the degree of site fidelity versus visual enhancement.

Figure 9. Gemini Image Generation for Site 3: (a) Original Site Image; (b) Gemini Generated Photo by Gemini Prompt; (c) Gemini by Generated Photo ChatGPT Prompt; (d) Gemini Generated Photo by Universal Prompter.



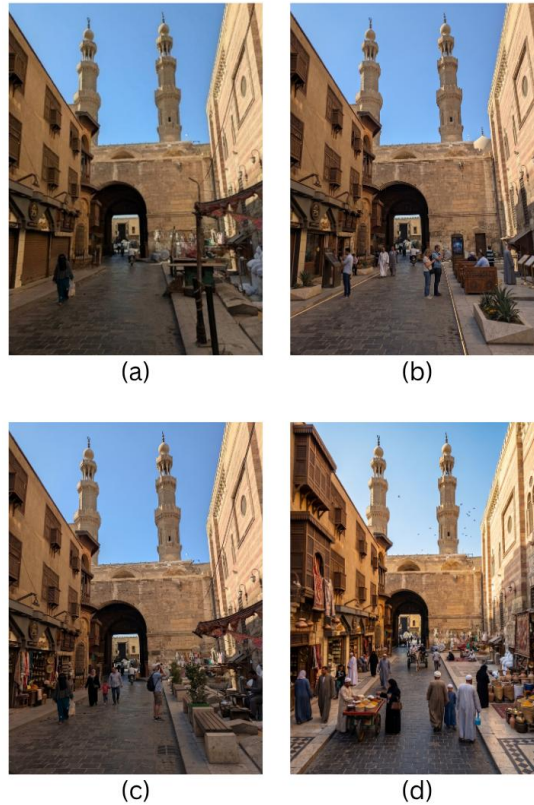
Figure 10 evaluates three Gemini-generated visualization against the original photograph of Site 4. Image (b), produced using the Gemini prompt, enhances the heritage setting through improved paving, pedestrian organization, and streetscape refinement while preserving the dominant minaret and dome. Image (c), generated using the ChatGPT prompt, remains closest to the original architectural composition with relatively limited intervention. Image (d), generated using the Universal Prompter, creates a more active and visually enhanced streetscape but introduces greater reinterpretation of the surrounding urban context. Overall, the outputs differ mainly in the balance between preservation fidelity and public-realm enhancement.

Figure 10. Gemini Image Generation for Site 4: (a) Original Site Image; (b) Gemini Generated Photo by Gemini Prompt; (c) Gemini Generated Photo by ChatGPT Prompt; (d) Gemini Generated Photo by Universal Prompter.



As shown in Figure 11 for site 5, image (b) produced using the Gemini prompt, preserves the main arched passage, twin minarets, and narrow street character while introducing a cleaner and more organized public realm. Image (c), generated using the ChatGPT prompt, remains closest to the original site configuration, with limited intervention and relatively strong preservation of the existing spatial character. Image (d), generated using the Universal Prompter, creates a more active and visually enhanced heritage streetscape, but it introduces greater reinterpretation of the surrounding commercial edges and overall urban setting. Overall, the three outputs differ mainly in their balance between preservation fidelity and streetscape enhancement.

Figure 11. Gemini Image Generation for Site 5: (a) Original Site Image; (b) Gemini Generated Photo by Gemini Prompt; (c) Gemini Generated Photo by ChatGPT Prompt; (d) Gemini Generated Photo by Universal Prompter.



4.3. Evaluation Performance

To systematically evaluate and compare the performance of Gemini Pro 3.1 Pro and ChatGPT 5.6 Sol across the generated visual scenarios, a structured multi-dimensional evaluation matrix was established. The assessment framework spans five core criteria categories, encompassing specific metrics assigned with relative weighting percentages to reflect architectural and visual priorities as shown in Table 1, given the likert scale is used (1= very weak, while 5=excellent).

Table 1. The Evaluation Metric

Criteria Category	Specific Metric	Evaluation Description
1. Prompt Adherence & Control	“Alter” Constraint Execution (10%)	How accurately the model executes specific physical or visual changes instructed in the prompt (e.g., repairing damaged masonry or adding a specific contemporary element)(Tan et al., 2026).
	“Retain” Constraint Execution (10%)	The model's ability to strictly preserve designated historical elements (e.g., original layout, proportions, material character, and architectural style)(Tan et al., 2026).
2. Aesthetic Compatibility & Visual Quality	Material Expression & Surface Texture (10%)	The realistic rendering of physical materials, particularly historic masonry, rough limestone, and aged brick (Tan et al., 2026).
	Color Harmony & Contrast (10%)	The accuracy, richness, and harmony of the colour palettes applied to the heritage structure and its environment (Tan et al., 2026)
	Lighting & Shadow Quality (10%)	The effective simulation of realistic natural lighting, soft shadows, and atmospheric depth (Tan et al., 2026).

	Composition & Spatial Arrangement (10%)	The maintenance of realistic architectural geometry, structural logic, and proportional harmony (Tan et al., 2026).
3. Ethical Integrity & Authenticity	Historical Reliability (No Hallucinations) (10%)	The extent to which the AI avoids inventing speculative architectural elements, distorting historical truth, or introducing regional/stylistic biases (Osman, 2018; Tan et al., 2026).
	Minimal Intervention Principle (10%)	The model's respect for the conservation principle that new additions should be visually subordinate and materially compatible with the original structure (Osman, 2018).
4. Functional Performance & Human Interaction	Human Interaction & Activity Visibility (5%)	The model's capacity to realistically embed human figures to communicate scale, liveliness, and functional occupation of the space (Tan et al., 2026)
	Adaptability to Contemporary Needs (5%)	How effectively the AI visualizes interior or exterior reconfigurations (adaptive reuse) to accommodate new functions (Tan et al., 2026).
5. Contextual Integration	Contextual & Landscape Compatibility (10%)	The ability of the AI to seamlessly blend the preserved heritage structure with the surrounding urban skyline and landscape (Tan et al., 2026).

For Site 1 in Al-Mu'izz Street, the evaluation showed a clear advantage for Gemini 3.1 Pro as the image generator. Gemini generated with either the Gemini-derived or ChatGPT-derived prompt achieved the joint-highest weighted score of 89%, followed by ChatGPT generated with the Gemini prompt at 85% and Gemini generated with the Universal Prompter at 80%. The lowest scores were obtained by ChatGPT using its own prompt (73%) and the Universal Prompter (60%). The differences were mainly associated with preservation-sensitive criteria. Gemini generated with the ChatGPT prompt received the maximum score of 5 for retain-constraint execution, composition and spatial arrangement, historical reliability, minimal intervention, and contextual compatibility, indicating strong preservation of the site's architectural identity. In contrast, ChatGPT generated using its own prompt scored only 2 for retention, historical reliability, and minimal intervention, while ChatGPT using the Universal Prompter scored 1 in these three criteria. Nevertheless, ChatGPT showed strong visual qualities, particularly in material expression and lighting, where its own prompting configuration received scores of 5. Thus, Site 1 demonstrates that a visually refined output did not necessarily correspond to stronger historical fidelity, and that Gemini provided a more balanced preservation-oriented result for this site, as shown in Figure 12.

Figure 12. Evaluation Result for Site 1.

Evaluating Generative AI Models for Historic Urban Heritage Visualization in Smart Cities

Specific Metric	Likert scale					
	ChatGPT 5.6 Sol generated by ChatGPT Prompts	ChatGPT 5.6 Sol generated by Gemini Prompts	ChatGPT 5.6 Sol generated by Universal Prompter	Gemini 3.1 Pro generated by Gemini Prompts	Gemini 3.1 Pro generated by ChatGPT Prompts	Gemini 3.1 Pro generated by Universal Prompter
"Alter" Constraint Execution 10%	4	4	4	4	4	4
"Retain" Constraint Execution 10%	2	4	1	4	5	4
Material Expression & Surface Texture 10%	5	4	4	5	4	4
Colour Harmony & Contrast 10%	4	4	4	4	4	4
Lighting & Shadow Quality 10%	5	4	4	5	4	4
Composition & Spatial Arrangement 10%	4	5	4	5	5	4
Historical Reliability (No Hallucinations) 10%	2	4	1	4	5	4
Minimal Intervention Principle 10%	2	4	1	4	5	4
Human Interaction & Activity Visibility 5%	4	5	4	4	3	4
Adaptability to Contemporary Needs 5%	5	4	4	5	4	4
Contextual & Landscape Compatibility 10%	4	5	3	5	5	4

For Site 2, the highest overall performance was achieved by Gemini generated using the Gemini prompt (93%), followed closely by ChatGPT generated using the Gemini prompt (92%) and Gemini generated using the ChatGPT prompt (91%). ChatGPT using its own prompt achieved 89%, whereas both Universal Prompter configurations obtained considerably lower scores of 66%. The strongest Gemini-specific configuration achieved maximum scores for retention, material expression, lighting, composition, historical reliability, adaptability, and contextual compatibility. Gemini generated with the ChatGPT prompt was particularly strong in conservation-related criteria, receiving 5 for retention, composition, historical reliability, minimal intervention, human activity, and contextual compatibility. The Universal Prompter produced the clearest weakness at this site: regardless of the generator, retention and historical reliability fell to 2, and contextual compatibility was limited to 3. Figure 13 illustrates Site 2 therefore reinforces the advantage of AI-generated, task-specific prompts over the generalized universal approach, while also showing that both Gemini and ChatGPT were capable of high performance when guided by appropriately structured prompts.

Figure 13. Evaluation Result for Site 2.

Evaluating Generative AI Models for Historic Urban Heritage Visualization in Smart Cities

Specific Metric	Likert scale					
	ChatGPT 5.6 Sol generated by ChatGPT Prompts	ChatGPT 5.6 Sol generated by Gemini Prompts	ChatGPT 5.6 Sol generated by Universal Prompter	Gemini 3.1 Pro generated by Gemini Prompts	Gemini 3.1 Pro generated by ChatGPT Prompts	Gemini 3.1 Pro generated by Universal Prompter
"Alter" Constraint Execution 10%	4	4	4	4	4	4
"Retain" Constraint Execution 10%	4	5	2	5	5	2
Material Expression & Surface Texture 10%	5	5	4	5	4	4
Colour Harmony & Contrast 10%	4	4	4	4	4	4
Lighting & Shadow Quality 10%	5	5	4	5	4	4
Composition & Spatial Arrangement 10%	5	5	4	5	5	4
Historical Reliability (No Hallucinations) 10%	4	4	2	5	5	2
Minimal Intervention Principle 10%	4	4	2	4	5	2
Human Interaction & Activity Visibility 5%	4	5	4	4	5	4
Adaptability to Contemporary Needs 5%	5	5	4	5	4	4
Contextual & Landscape Compatibility 10%	5	5	3	5	5	3

The ranking changed for Site 3, demonstrating that performance was influenced by the characteristics of the individual heritage setting. Gemini generated using the ChatGPT prompt achieved the highest score of 91%, followed by ChatGPT using its own prompt at 89%. Gemini generated using the Universal Prompter reached 81%, while both Gemini using its own prompt and ChatGPT using the Gemini prompt obtained 79%, and ChatGPT using the Universal Prompter scored 77%. The Gemini–ChatGPT prompt combination was particularly strong in preservation fidelity, receiving 5 for retention, composition, historical reliability, minimal intervention, human interaction, and contextual compatibility. By contrast, Gemini using its own prompt and ChatGPT using the Gemini prompt showed weaker preservation scores, particularly for retention, historical reliability, and minimal intervention, despite maintaining strong material and lighting performance. The Universal Prompter performed comparatively better at Site 3 than at Sites 1 and 2, particularly in human interaction, but remained less consistent in conservation-oriented criteria. Figure 14 results indicate that the ChatGPT-derived prompt was especially effective for maintaining the spatial and heritage character of Site 3 when implemented by Gemini.

Figure 14. Evaluation Result for Site 3.

Evaluating Generative AI Models for Historic Urban Heritage Visualization in Smart Cities

Specific Metric	Likert scale					
	ChatGPT 5.6 Sol generated by ChatGPT Prompts	ChatGPT 5.6 Sol generated by Gemini Prompts	ChatGPT 5.6 Sol generated by Universal Prompter	Gemini 3.1 Pro generated by Gemini Prompts	Gemini 3.1 Pro generated by ChatGPT Prompts	Gemini 3.1 Pro generated by Universal Prompter
"Alter" Constraint Execution 10%	4	4	4	4	4	4
"Retain" Constraint Execution 10%	4	3	4	3	5	4
Material Expression & Surface Texture 10%	5	5	4	5	4	4
Colour Harmony & Contrast 10%	4	4	4	4	4	4
Lighting & Shadow Quality 10%	5	5	4	5	4	4
Composition & Spatial Arrangement 10%	5	4	4	4	5	4
Historical Reliability (No Hallucinations) 10%	4	3	3	3	5	4
Minimal Intervention Principle 10%	4	3	3	3	5	4
Human Interaction & Activity Visibility 5%	4	4	5	4	5	5
Adaptability to Contemporary Needs 5%	5	5	4	5	4	4
Contextual & Landscape Compatibility 10%	5	4	4	4	5	4

Site 4 evaluation was generally high across most configurations, with ChatGPT generated using the Gemini prompt producing the highest individual score in the entire five-site evaluation at 94%. Gemini using its own prompt and ChatGPT using the Universal Prompter both achieved 91%, followed by ChatGPT using its own prompt at 89%, Gemini using the ChatGPT prompt at 88%, and Gemini using the Universal Prompter at 80%. ChatGPT generated with the Gemini prompt achieved maximum scores for retention, material expression, colour harmony, lighting, composition, human activity, adaptability, and contextual compatibility, indicating a particularly strong combination of visual quality and preservation. Gemini generated with the ChatGPT prompt performed strongly in retention, composition, historical reliability, minimal intervention, and contextual compatibility, but obtained lower scores for human interaction and adaptability. Site 4 is also notable because the ChatGPT Universal Prompter configuration performed unusually well, receiving scores of 5 for material, colour, lighting, composition, adaptability, and contextual compatibility. Nevertheless, this stronger site-specific result did not reflect the Universal Prompter's overall performance across the five sites, as shown in Figure 15.

Figure 15. Evaluation Result for Site 4.

Evaluating Generative AI Models for Historic Urban Heritage Visualization in Smart Cities

Specific Metric	Likert scale					
	ChatGPT 5.6 Sol generated by ChatGPT Prompts	ChatGPT 5.6 Sol generated by Gemini Prompts	ChatGPT 5.6 Sol generated by Universal Prompter	Gemini 3.1 Pro generated by Gemini Prompts	Gemini 3.1 Pro generated by ChatGPT Prompts	Gemini 3.1 Pro generated by Universal Prompter
"Alter" Constraint Execution 10%	4	4	4	4	4	4
"Retain" Constraint Execution 10%	4	5	4	4	5	4
Material Expression & Surface Texture 10%	5	5	5	5	4	4
Colour Harmony & Contrast 10%	4	5	5	5	4	4
Lighting & Shadow Quality 10%	5	5	5	5	4	4
Composition & Spatial Arrangement 10%	5	5	5	5	5	4
Historical Reliability (No Hallucinations) 10%	4	4	4	4	5	4
Minimal Intervention Principle 10%	4	4	4	4	5	3
Human Interaction & Activity Visibility 5%	4	5	4	4	3	5
Adaptability to Contemporary Needs 5%	5	5	5	5	3	5
Contextual & Landscape Compatibility 10%	5	5	5	5	5	4

ChatGPT generated using its own prompt achieved the highest score of 91%, followed by ChatGPT using the Gemini prompt and Gemini using the ChatGPT prompt, both at 90%, and Gemini using its own prompt at 89% for Site 5. A much larger reduction was observed for the Universal Prompter, with Gemini achieving 76% and ChatGPT 67%. ChatGPT using its own prompt performed particularly strongly in retention, material expression, lighting, composition, adaptability, and contextual compatibility, while Gemini using the ChatGPT prompt achieved maximum scores for retention, composition, historical reliability, human interaction, adaptability, and contextual compatibility. The Universal Prompter again showed its principal weaknesses in heritage-sensitive criteria. ChatGPT using the universal approach scored only 2 for retention, historical reliability, and minimal intervention, while the corresponding Gemini configuration obtained 3 for these criteria. Both universal configurations nevertheless retained relatively good performance in human interaction and contemporary adaptability. Figure 15 therefore confirms that the principal limitation of the Universal Prompter was not general image quality or activity representation, but the preservation of site-specific architectural identity and historical reliability.

Figure 15. Evaluation Result for Site 5.

Evaluating Generative AI Models for Historic Urban Heritage Visualization in Smart Cities

Specific Metric	Likert scale					
	ChatGPT 5.6 Sol generated by ChatGPT Prompts	ChatGPT 5.6 Sol generated by Gemini Prompts	ChatGPT 5.6 Sol generated by Universal Prompter	Gemini 3.1 Pro generated by Gemini Prompts	Gemini 3.1 Pro generated by ChatGPT Prompts	Gemini 3.1 Pro generated by Universal Prompter
"Alter" Constraint Execution 10%	4	4	4	4	4	4
"Retain" Constraint Execution 10%	5	4	2	5	5	3
Material Expression & Surface Texture 10%	5	5	4	4	4	4
Colour Harmony & Contrast 10%	4	5	4	4	4	4
Lighting & Shadow Quality 10%	5	5	4	4	4	4
Composition & Spatial Arrangement 10%	5	4	4	5	5	4
Historical Reliability (No Hallucinations) 10%	4	4	2	5	5	3
Minimal Intervention Principle 10%	4	4	2	4	4	3
Human Interaction & Activity Visibility 5%	4	5	5	4	5	5
Adaptability to Contemporary Needs 5%	5	5	4	5	5	
Contextual & Landscape Compatibility 10%	5	5	3	5	5	4

Across the five sites, Gemini 3.1 Pro demonstrated the stronger overall profile as an image generator for heritage-preservation applications. Averaged across all prompting conditions, Gemini achieved approximately 84.9%, compared with 82.1% for ChatGPT 5.6 Sol. When only the two AI-generated prompting approaches were considered, excluding the weaker Universal Prompter, the difference was 89.0% for Gemini and 87.1% for ChatGPT. The main strength of Gemini was its conservation fidelity: its average scores for retain-constraint execution and historical reliability were both 4.20/5, compared with 3.53/5 and 3.27/5, respectively, for ChatGPT, while it also performed better in minimal intervention (3.93 vs. 3.27) and slightly better in contextual compatibility (4.53 vs. 4.40). Gemini therefore showed a greater ability to preserve original architectural features and spatial relationships while limiting unsupported alterations. ChatGPT's main strengths were instead associated with visual and representational quality, achieving higher averages for material expression and surface texture (4.67 vs. 4.27), lighting and shadow quality (4.67 vs. 4.27), human interaction (4.40 vs. 4.27), and adaptability to contemporary needs (4.67 vs. 4.47). Its principal weakness for heritage-preservation visualization was therefore not image realism, but a greater tendency to reinterpret the existing architecture, which reduced historical reliability and adherence to the minimal-intervention principle.

The comparison of the systems as prompt generators produced a different pattern. When the resulting images from both generators were combined, Gemini-derived prompts averaged 88.1% and ChatGPT-derived prompts 88.0%, indicating no meaningful overall difference in their total effectiveness. Their strengths, however, were distinct. ChatGPT-derived prompts produced stronger preservation-oriented results, averaging 4.40/5 for retention, 4.30/5 for historical reliability, 4.20/5 for minimal intervention, 4.90/5 for composition, and 4.90/5 for contextual compatibility. Gemini-derived prompts were stronger in visual and functional enhancement, achieving 4.80/5 for material expression, 4.80/5 for lighting, 4.30/5 for colour harmony, 4.40/5 for human interaction, and 4.90/5 for adaptability to contemporary needs. The strongest complete generator-prompt combination was Gemini generating images from ChatGPT-derived prompts, with an overall mean of 89.8%, followed by Gemini using its own prompts (88.2%) and ChatGPT using Gemini-derived prompts (88.0%). This cross-model performance suggests that the strengths of the two systems were complementary: ChatGPT was particularly effective at formulating preservation-oriented instructions, while Gemini was more reliable at translating such instructions into historically faithful images; conversely, Gemini prompts supported

stronger visual enhancement when executed by ChatGPT. The Universal Prompter was substantially weaker overall at 74.4%, particularly in retention, historical reliability, minimal intervention, and contextual compatibility, indicating that a generalized prompt was less suitable than AI-specific prompting for site-sensitive heritage visualization.

5. Conclusion

This study set out to evaluate the efficacy and accuracy of generative AI tools in visualizing historical urban heritage, utilizing Al-Mu'izz Street in Cairo as a primary case study to establish a foundational framework for AI-assisted urban preservation in smart cities. Through a rigorous comparative assessment of generative models and prompt engineering strategies, the investigation successfully addressed its core objectives:

1. **Assessing AI Capabilities:** The findings demonstrate that modern generative models are capable of visualizing complex urban heritage scenarios, achieving overall accuracy profiles of 84.9% for Gemini 3.1 Pro and 82.1% for ChatGPT 5.6 Sol across evaluation sites. However, model performance varied significantly by design dimension; Gemini excelled in conservation fidelity, retention, and historical reliability (averaging 4.20/5), whereas ChatGPT prioritized visual presentation, material expression, and lighting quality (averaging 4.67/5).
2. **Optimizing Prompt Engineering:** The results confirm that generalized instructions, such as the Universal Prompter (scoring 74.4%), are inadequate for site-sensitive heritage visualization. Instead, tailored cross-model prompt structuring yielded superior results. Specifically, the combination of Gemini generating images from ChatGPT-derived prompts achieved the highest overall performance (89.8%), proving that ChatGPT is exceptionally effective at formulating preservation-oriented textual instructions, which Gemini then translates into historically faithful visuals.
3. **Benchmarking and Validation:** By benchmarking AI outputs against architectural documentation and historical layouts, the study established that maintaining adherence to the minimal-intervention principle remains a challenge for aesthetic-driven models. While models easily enhance surface textures and lighting, unchecked generation risks over-interpreting historic fabrics and compromising authenticity.
4. **Smart City and Digital Twin Integration:** Ultimately, this research provides a scalable framework for urban planners and heritage conservators looking to embed generative AI into digital twin ecosystems and participatory urban planning workflows.

In conclusion, generative AI serves as a powerful accelerator for early-stage design exploration and heritage visualization, but its utility in smart cities depends heavily on hybrid workflows—pairing precision prompt strategies with strict conservation criteria to ensure that digital reconstructions remain culturally authentic, historically reliable, and publicly legitimate.

4. Acknowledgements

The authors would like to acknowledge the support provided by the Collage of Design and Built Environment at King Fahd University of Petroleum and Minerals (KFUPM).

References

- Abdelalim, M. (2025). Heritage Preservation Using Laser Scanning: Architectural Digital Twins Using Al-Mu'izz Street as a Case Study. *Buildings*, 15(9), 1480.
<https://doi.org/10.3390/buildings15091480>
- Abdel-Hadi, A., El-Nachar, E., & Safiieldin, H. (2009). Pedestrian Street Life in Historic Cairo. AI in architectural concept design: Reimagining the conflict-affected areas in Syria. (n.d.). Al-Mu'izz Street – The Heart of Islamic Cairo. (2026, July 7). Monkey's Tale.
<https://monkeystale.ca/2026/07/07/al-muizz-street-the-heart-of-islamic-cairo/>

- Angelidou, M., Karachaliou, E., Angelidou, T., & Stylianidis, E. (2017). CULTURAL HERITAGE IN SMART CITY ENVIRONMENTS. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W5, 27–32. <https://doi.org/10.5194/isprs-archives-XLII-2-W5-27-2017>
- Artchitectours. (2025, June 26). Artificial intelligence and heritage conservation ► the latest insights. Artchitectours. <https://www.artchitectours.com/artificial-intelligence-heritage-conservation/>
- Borda, A., & Bowen, J. P. (2017). Smart Cities and Cultural Heritage – A Review of Developments and Future Opportunities. *Electronic Visualisation and the Arts (EVA 2017)*. <https://doi.org/10.14236/ewic/EVA2017.2>
- ElKatsha, M. F. (2000). The Evolution Of Al-Azhar Street Al-Qahira, Egypt.
- Ghaith, K. (2024). AI Integration in Cultural Heritage Conservation – Ethical Considerations and the Human Imperative. *International Journal of Emerging and Disruptive Innovation in Education : VISIONARIUM*, 2(1). <https://doi.org/10.62608/2831-3550.1022>
- Lee, E.-J., & Park, S.-J. (2025). Generative AI for Biophilic Design in Historic Urban Alleys: Balancing Place Identity and Biophilic Strategies in Urban Regeneration. *Land*, 14(10), 2085. <https://doi.org/10.3390/land14102085>
- Moez Street in Cairo: An Ultimate Local's Guide. (2024, March 15). <https://vanillatravels.net/moez-street-in-cairo/>
- Mukhamedjanova, S. (2025). The Use of Artificial Intelligence in Architectural Restoration: A Review. *Proceedings of the 9th International Conference on Future Networks and Distributed Systems*, 1179–1185. <https://doi.org/10.1145/3789692.3789849>
- Osman, K. A.-A. (2018). Heritage conservation management in Egypt. *Ain Shams Engineering Journal*, 9(4), 2907–2916. <https://doi.org/10.1016/j.asej.2018.10.002>
- Ran, M., Lin, X., Tian, M., Cong, J., Chen, S., Wang, T., & Jiang, Y. (2025). Generative AI-Based Night Scene Design of Historic Districts: Value Transition from Basic Lighting to Cultural Narrative. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-M-9-2025, 1273–1280. <https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-1273-2025>
- Sektani, H. H. J., Miran, F. D., & Abdullah, H. K. (2026). AI-Assisted Visualisation of Heritage Conservation Interventions: An Exploratory Study of Community Preferences. *Heritage*, 9(6), 226. <https://doi.org/10.3390/heritage9060226>
- Streets of Islamic Cairo: A configuration of urban themes and patterns. (n.d.).
- Tan, X., Xie, H., Xia, C., Ying, Y., Hokoi, S., & Li, Y. (2026). Artificial intelligence in five key phases of architectural heritage conservation: A systematic review. *Journal of Building Engineering*, 117, 114862. <https://doi.org/10.1016/j.jobbe.2025.114862>