JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 27, No. 10, pp. 2729-2740, Oct. 2026

Check for updates

생성형 AI 기반 한국 전통건축 이미지의 재현 특성과 활용 적합성 연구

김 지 은[1] · 박 진 완[2*]
[1]중앙대학교 첨단영상대학원 영상학과 박사과정
[2]중앙대학교 첨단영상대학원 영상학과 교수

Representational Characteristics and Intended-Use Suitability of Generative AI-Based Images of Korean Traditional Architecture

Ji-Eun Kim[1] · Jin Wan Park[2*]

[1]Doctoral Course, Graduate School of Advanced Imaging, Chung-Ang University, Seoul 06974, Korea
[2]Professor, Graduate School of Advanced Imaging, Chung-Ang University, Seoul 06974, Korea

[요 약]

생성형 인공지능은 문화유산 콘텐츠 제작에 활발히 활용되고 있으나, 전통 건축 이미지 생성 과정에서 구조적·역사적 오류가 빈번하게 발생한다. 본 연구는 생성형 AI가 생성한 한국 전통 건축 이미지의 활용 목적별 검증 수준을 분석하였다. Midjourney와 Gemini를 활용하여 4개 전통 건축물을 대상으로 이미지를 생성하고, 정량분석, 전문가 검증 및 활용성 평가를 실시하였다. 분석 결과, 생성형 AI는 전통 건축의 시각적 정체성을 효과적으로 재현하였으나, 공포 구조, 결구 방식, 현판 문자, 단청 문양 등 핵심 요소에서는 반복적인 오류가 확인되었다. 생성 이미지는 창작 및 홍보 목적에서는 높은 활용 가능성을 보였으나, 교육 및 역사 고증 목적에서는 보다 엄격한 검수가 요구되었다. 이에 본 연구는 활용 목적 기반의 문화유산 이미지 평가 체계인 Heritage AI Verification Framework를 제안하고, 문화유산 콘텐츠 제작을 위한 실무적 검증 지침을 제시하였다.

[Abstract]

Generative artificial intelligence (AI) is increasingly being used in cultural heritage content production. However, images of traditional architecture often contain structural and historical inaccuracies. This study examines how the intended use of AI-generated images of Korean traditional architecture affect verification requirements. The images of four representative heritage buildings generated using Midjourney and Gemini were subjected to quantitative analysis, expert review, and usability assessments. Results reveal that AI effectively reproduces traditional architecture's visual identity but frequently introduces errors in bracket structures, joinery systems, signboard inscriptions, and decorative painting patterns. Although the images had high usability for creative and promotional purposes, their educational and historical applications required stricter verification. This study proposes the development of the Heritage AI Verification Framework to provide practical guidelines and decision-making criteria for AI-generated cultural heritage imagery evaluation based on intended use.

색인어 : 문화유산 콘텐츠, 생성형 인공지능, 이미지 검증, 한국 전통 건축, 활용 적합성

Keyword : Cultural Heritage Content, Generative AI, Image Verification, Korean Traditional Architecture, Intended Use Suitability

http://dx.doi.org/10.9728/dcs.2026.27.10.2729

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-CommercialLicense(http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 14 July 2026; **Revised** 12 August 2026
Accepted 14 September 2026

***Corresponding Author; Jin Wan Park**

Tel: +82-2-820-5710
E-mail: jinpark@cau.ac.kr

Ⅰ. 연구의 배경 및 목적

생성형 인공지능은 이미지 제작의 속도와 접근성을 향상시키며 문화유산 시각 콘텐츠의 생산 방식에 빠른 변화를 가져오고 있다. 한국 전통 건축은 관광 홍보, 교육 자료, 전시 콘텐츠, 게임 및 영상 배경 등에서 활용되는 시각 자원으로 주목받고 있으나, 생성형 AI가 제작한 전통 건축 이미지는 겉보기에 그럴듯하더라도 건축 구조, 장식 요소, 재료 표현에서 왜곡을 포함하는 경우가 적지 않다. 이러한 왜곡은 창작·홍보 목적에서는 수용될 수 있으나 교육·학술·고증 분야에서는 콘텐츠의 신뢰성을 저해할 수 있다. 문제는 이러한 생성 이미지가 검증 기준이 부재한 상태로 이미 현장에서 광범위하게 유통되고 있어, 콘텐츠 디자이너와 실무자가 채택 여부를 체계적 근거 없이 판단해야 한다는 점이다.

기존 연구는 생성형 AI의 표현 성능과 공간 재현 능력을 중심으로 평가해 왔다. Ibrahim 등[1]은 주요 생성 모델의 전통 건축 이미지 생성 성능을 비교하였고, Odiah와 Gosling[2]은 건축 공간의 분위기 전달력을 인간 평가로 검증하였으며, Paananen 등[3]은 설계 초기 단계의 아이디어 발산 지원 가능성을 보고하였다. 그러나 문화유산 분야에서 AI 활용의 가능성과 한계가 논의되고 있음에도[4], 활용 목적에 따라 요구되는 검증 수준과 허용 가능한 왜곡 범위를 체계적으로 제시한 연구는 부족하며, Pansoni 등[5]의 AI 윤리 프레임워크 역시 거버넌스 차원에 집중되어 실무적 검증 가이드라인은 다루지 못하였다.

이에 본 연구는 생성형 AI가 생성한 한국 전통 건축 이미지의 시각적 변형 양상을 분석하고, 활용 목적별로 요구되는 정밀도와 검수 항목을 체계화하는 것을 목적으로 한다. 이를 위해 정량 유사성 분석, 전통 건축 전문가 검증, 콘텐츠 실무자·전공자의 활용성 평가를 수행하고, 각 분석 결과를 삼각검증하여 Heritage AI Verification Framework를 제안하고자 한다.

이를 바탕으로 본 연구는 다음의 연구 질문을 설정하였다. (RQ1) 생성형 AI가 한국 전통 건축을 재현하는 과정에서 반복적으로 나타나는 왜곡은 어떤 구조·장식 요소에 집중되는가? (RQ2) 생성 조건과 이미지 범위에 따라 재현 수준은 어떻게 달라지는가? (RQ3) 생성 이미지의 활용 적합성은 활용 목적에 따라 어떻게 달라지며, 평가 주체에 따라 어떠한 차이를 보이는가? (RQ4) 이상의 결과를 바탕으로 활용 목적별 검증 기준을 어떻게 체계화할 수 있는가? 각 연구 질문은 RQ1은 4-2절, RQ2는 4-1절, RQ3은 4-3·4-4절, RQ4는 5-2절에서 다루었다.

Ⅱ. 이론적 배경

2-1 생성형 AI 문화유산 이미지와 활용 목적별 요구 정밀도

생성형 인공지능은 문화유산 분야에서도 활용 범위를 빠르게 확대하고 있다. Stable Diffusion, Midjourney, DALL·E, Gemini 등의 모델은 건축물, 유적지, 전통 공예품을 시각적으로 재구성하는 데 활용되며, 관광 홍보, 전시, 교육, 디지털 아카이브 등 다양한 분야에서 적용되고 있다. 그러나 문화유산 이미지는 특정 시대와 지역의 역사·문화 정보를 담고 있으므로 일반 이미지보다 높은 수준의 신뢰성과 정확성이 요구된다.

선행연구에 따르면 생성형 AI는 문화유산 재현 과정에서 구조적·역사적 왜곡을 반복적으로 발생시키는 것으로 나타났다. Sukkar 등[6]은 Midjourney가 이슬람 건축 이미지를 생성하는 과정에서 서로 다른 시대와 지역의 양식을 혼합하고 세부 건축 요소를 부정확하게 재현하는 경향을 보고하였다. 또한 청나라 궁궐 내부를 대상으로 한 연구[7]에서는 공간 깊이 과장, 비례 왜곡, 천장 요소 축소 등 6.5~40.6% 수준의 구조 오류가 관찰되었으며, 생성형 AI가 구조적 정확성보다 시각적 인상을 우선하는 경향이 있음이 지적되었다. 이러한 결과는 생성형 AI가 문화유산의 전체적 인상은 재현하더라도 핵심 구조 요소에서는 체계적 왜곡을 발생시킬 수 있음을 보여준다.

그러나 기존 연구는 주로 이미지 품질, 텍스트 정합성, 생성 성능과 같은 기술적 지표에 초점을 두어 왔으며, 생성 결과물이 실제 콘텐츠 제작 과정에서 어느 수준까지 활용 가능한지에 대한 논의는 상대적으로 부족하였다. 문화유산 이미지는 활용 목적에 따라 요구되는 정확성이 다르며, 동일한 이미지라도 창작, 홍보, 교육, 고증의 맥락에서 서로 다른 평가를 받을 수 있다. 본 연구에서 창작은 게임 컨셉아트, 영상 배경 등 예술적 변형이 허용되는 콘텐츠를, 홍보는 관광 안내물, 문화유산 홍보물 등 대상의 대표성이 요구되는 콘텐츠를, 교육은 교재, 해설 자료 등 정보의 정확성이 요구되는 콘텐츠를, 고증은 문화재 기록, 학술 자료 등 역사적 진정성이 요구되는 콘텐츠를 의미한다. 예를 들어 동일한 궁궐 이미지라도 게임 배경으로는 충분히 활용 가능하지만, 문화재 해설 자료로는 부적합할 수 있다. 따라서 생성형 AI 문화유산 이미지는 단일한 품질 기준이 아니라 활용 목적에 따른 요구 정밀도(level of required precision) 관점에서 검토될 필요가 있다.

2-2 왜곡 유형과 활용 목적 기반 검증의 필요성

생성형 AI를 활용한 시각적 재현 시도는 문화유산과 건축 분야 전반으로 확산되고 있다. Akhtar 등[8]은 하라파 문명의 소실된 안료를 재구성하기 위해 생성형 AI 기반의 디지털 색채 복원을 시도하였으며, Tan과 Luhrs[9]는 Midjourney를 건축 디자인 프로세스에 적용하여 초기 단계의 아이디어 발산에는 유용하나 구조적 정확성이 요구되는 단계에서는 한계가 있음을 보고하였다. Ibrahim 등[1] 역시 전통 건축 이미지 생성 연구에서 시각적 완성도와 건축적 정확성 사이에 유의미한 간극이 존재함을 확인하였다. 이처럼 활용이 확대되고 그 한계가 반복적으로 확인됨에 따라, 산출물의 유효성을

평가하기 위한 비판적 검증 기준의 필요성이 인접 분야에서 제기되고 있다. Gattiglia[10]는 고고학 분야의 AI 도입을 다루며, 기존 데이터의 편향이 재생산되거나 잠정적 해석이 확정적 사실로 고착될 위험을 지적하고 비판적 검증 체계의 시급성을 강조하였다.

이상의 선행연구는 연구 대상과 접근 방식에 따라 세 갈래로 구분된다. 이를 정리하고 본 연구의 위치를 함께 제시하면 표 1과 같다. 선행연구는 생성 성능과 건축적 활용 가능성, 문화유산 재현 과정에서의 오류, 그리고 AI 활용의 윤리적·거버넌스적 문제를 각각 다루어 왔다. 그러나 개별 생성 이미지에서 확인되는 왜곡 유형을 실제 콘텐츠 활용 목적과 연결하여 채택 가능 여부를 판단하는 연구는 상대적으로 제한적이다. 이에 본 연구는 한국 전통 건축을 대상으로 생성 이미지의 구조적·장식적 왜곡을 유형화하고, 전문가 검증과 활용성 평가를 결합하여 활용 목적별 검증 기준을 제안하고자 한다.

검증 기준을 수립하기 위해서는 먼저 무엇을 검증할 것인지, 즉 생성 이미지에서 어떤 왜곡이 발생하는지가 규명되어야 한다. 특히 전통 건축 이미지는 외형적 유사성을 넘어 엄격한 건축 구조와 역사적 진정성이 담보되어야 하므로, 검증의 대상이 되는 요소 또한 일반 이미지와 다르다. 공포, 결구, 활주와 같은 구조 요소와 단청, 현판과 같은 장식 요소는 건

표 1. 선행연구 유형별 비교 및 본 연구의 위치
Table 1. Comparison of prior research streams and the positioning of this study

Research Stream	Studies / Object & Method / Limitation & Contribution
Generative AI in architectural representation	Studies [1], [2], [3], [9] Object & Method: Benchmarking of text-to-image models; human rating of conveyed ambiance; ideation support in early design. Limitation / Contribution: Evaluation centered on generation performance and design ideation; heritage-specific accuracy criteria are absent.
AI reconstruction of cultural heritage	Studies [6], [7], [8] Object & Method:Style-mixing and detail errors in Islamic heritage; structural errors in Qing palace interiors; digital pigment restoration. Limitation / Contribution:Errors are documented as model limitations, but are not linked to the admissible scope of use.
AI ethics & critical verification frameworks	Studies [4], [5], [10] Object & Method:R&D agendas for digital heritage; ethical framework design; critical review of AI adoption in archaeology. Limitation / Contribution:Operates primarily at the governance and principle level and is not directly actionable for image-level adoption decisions by practitioners.
This Study	Object & Method:Korean traditional architecture; 48 images (4 sites × 2 scales × 3 conditions × 2 models); triangulation of quantitative indices, expert review (W1–W6), and usability ratings by 20 practitioners. Contribution: Maps distortion types to four intended uses through use-precision matching, yielding an operational verification checklist.

표 2. 한국 전통 건축 주요 구조 용어
Table 2. Key structural terms in Korean traditional architecture

Term	Description	Photo
Bracket Set (Gongpo)	Interlocking bracket assembly on top of columns that distributes roof load; its form varies by style and serves as a key indicator of period and formality	
Joinery System (Gyeolgu)	Traditional wooden joinery technique connecting members without nails; composed of interlocking joints such as mortise-and-tenon and dovetail	
Diagonal Brace (Hwalju)	Auxiliary column supporting eave corners or floor edges; a structural member preventing eaves from sagging	
Dancheong	Decorative polychrome painting on wooden members based on five cardinal colors; reflects the formality and function of the building	
Signboard (Hyeonpan)	Wooden plaque inscribed with the building's name in Chinese characters; its calligraphy and placement define identity, requiring accurate reproduction for historical documentation	
Stone Platform (Gidan)	Stone foundation platform of a building; composed of terraces, carved central stairway panels, and balustrades, determining hierarchy and spatial context	
Rafter (Seokkaerae)	Round timber members forming the roof frame; arrayed on purlins, they determine eave curvature and roof proportion	
Tall Pillar (Goju)	Tall interior column supporting the main crossbeam; creates height variation in the roof structure	

*Chinese characters are retained in their original form and are not relevant to the interpretation of this study

표 3. 생성형 AI 전통건축 이미지의 왜곡 유형
Table 3. Types of distortion in traditional architectural images generated by AI

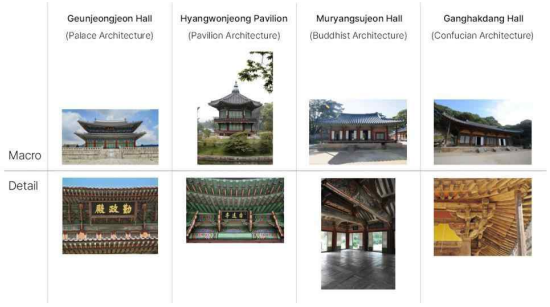
Style Blending (W1)	Ambiguous mixing of architectural styles from different periods or regions
Proportional Distortion (W2)	Mechanical and visual proportions among building members such as roofs, columns, and bracket sets are inconsistent with actual architecture
Structural Simplification (W3)	Omission of the complex joinery methods and framework structures characteristic of traditional architecture
Pattern Hallucination (W4)	Generation of fictional patterns or unnatural forms without historical basis
Material Misinterpretation (W5)	Distorted or inaccurate representation of textures and materials such as wood, stone, and roof tiles
Contextual Absence (W6)	Surrounding environment or topographical context of the building contradicts historical records

축물의 시대와 격식을 판별하는 근거이자 생성형 AI가 왜곡을 일으키기 쉬운 부위이기도 하다. 이에 본 연구는 선행연구와 예비분석을 바탕으로, 검증 시 주의 깊게 확인해야 할 전통 건축의 핵심 요소를 표 2에 정리하고, 생성 이미지에서 반복적으로 나타나는 왜곡을 여섯 가지 유형으로 정립하여 표 3에 제시하였다. 이 여섯 가지 왜곡 유형(W1~W6)은 이후 전문가 검증의 평가 기준으로 활용된다.

Ⅲ. 연구 방법

3-1 연구 설계 및 분석 코퍼스

본 연구는 제한된 사례를 대상으로 생성형 AI의 한국 전통 건축 재현 특성과 왜곡 양상을 탐색하는 탐색적 연구(exploratory study)로 설계하였다. 따라서 본 연구의 결과는 한국 전통 건축 전반이나 모든 생성형 AI 모델에 대한 통계적 일반화보다는, 사례에서 반복적으로 확인된 오류 유형과 활용 단계별 검증 필요성을 제시하는 데 목적이 있다. 분석에는 정량적 자료와 질적 자료를 독립적으로 수집·분석한 후 통합하여 해석하는 혼합 연구 방법(mixed methods)을 적용하였으며[11],[12], 정량 유사성 분석, 전통 건축 전문가 검증, 콘텐츠 활용성 평가를 상호 비교하는 삼각 검증(triangulation) 설계를 채택하였다[13].

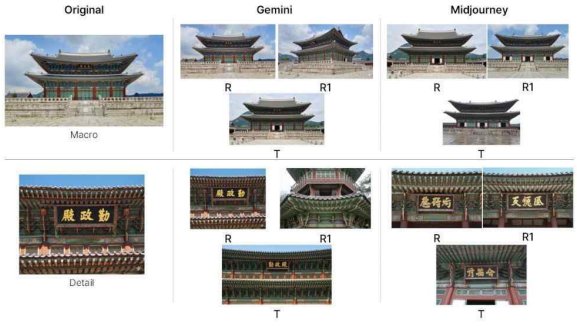


*Chinese characters are retained in their original form and are not relevant to the interpretation of this study

그림 1. 분석 대상 한국 전통건축물 원본 이미지(출처: 국가유산청 공공저작물)

Fig. 1. Original images of the selected Korean traditional architectural heritage sites (source: Korea Heritage Service Public Image Archive)

분석 대상은 통계적 대표성이 아니라 유형적 다양성의 확보를 우선하는 최대 변이 표집(maximum variation sampling)을 적용하여 선정하였다[14]. 이는 소수의 사례를 다루면서도 대상 간 차이가 큰 사례를 의도적으로 포함함으로써 사례 전반에 공통된 양상과 사례별로 상이한 양상을 함께 확인할 수 있게 하며, 특정 건축물의 재현 성능보다 전통 건축 전반에서 반복되는 왜곡 유형을 규명하고자 하는 본 연구의 목적에 부합한다. 연구 대상은 그림 1과 같이 궁궐 건축인 근정전, 불교 건축인 무량수전, 정자 건축인 향원정, 유교 건축인 강학당으로 구성하였다. 이들 건축물은 구조 체계, 장식 방식 및 공간 구성 원리가 상이하여 생성형 AI가 서로 다른 전통 건축 유형을 어떻게 재현하는지 비교·분석하기에 적합하다. 각 건축물은 전체 형태와 공간 구성을 확인할 수 있는 전체(Macro) 이미지와 공포, 단청, 결구 등 세부 구조를 관찰할 수 있는 디테일(Detail) 이미지로 구분하여 구성하였다.



*Chinese characters are retained in their original form and are not relevant to the interpretation of this study

그림 2. Gemini와 Midjourney의 생성 조건별 비교(근정전 예시)

Fig. 2. Comparison of generation conditions (T, R1, and R) using Gemini and Midjourney

이미지 생성 조건은 그림 2와 같이 세 유형으로 설정하였다. 텍스트 프롬프트만 사용하는 Text-to-Image(T), 참조 이미지를 함께 입력하여 기존 시점을 유지한 상태로 생성하는 Reference-Same View(R1), 참조 이미지에 시점 지시구를 부가하여 새로운 시점을 유도하는 Reference-Novel View(R)이다. 세 조건은 동일한 기저 프롬프트를 공유하며, 참조 이미지의 입력 여부와 시점 지시구의 부가 여부에서만 차이를 갖는다. 기저 프롬프트는 표 2에 정리한 전통 건축의 구조·장식 요소가 명시적으로 기술되도록 작성하였으며, 프롬프트 원문과 조건별 구성 방식은 부록에 제시하였다.

다만 R 조건은 기저 프롬프트의 기존 시점 서술을 유지한 상태에서 시점 지시구를 후행 부가하는 방식으로 구성되어, 일부 사례에서 두 시점 지시가 중첩 또는 충돌하였다. 따라서 본 연구의 R 조건은 시점변환 성능을 측정하기 위한 조건이 아니라, 참조 이미지를 기반으로 새로운 장면을 구성할 때 나타나는 재현 특성을 탐색하기 위한 조건으로 설정하였다. 상세 내역은 부록에 제시하였다.

생성 모델은 Gemini와 Midjourney를 사용하였다. 본 연구는 생성 모델의 성능 우열을 비교하는 연구가 아니므로 벤치마크 순위나 시장 점유율을 기준으로 모델을 선정하지 않고, 서로 다른 성격의 상용 생성 환경을 사례로 포함하고자 하였다. Midjourney는 건축 및 문화유산 이미지 생성 연구에서 반복적으로 분석 대상이 되어 온 상용 서비스로[1],[6],[9] 선행연구와의 비교 가능성을 고려하여 선정하였다. Gemini는 별도의 환경 구축 없이 대화형 인터페이스에서 참조 이미지 입력과 생성이 가능하여, 전문적 도구 사용 경험이 없는 콘텐츠 실무자도 접근할 수 있는 생성 환경을 대표한

다는 점을 고려하였다.

아울러 두 모델 모두 참조 이미지 기반 생성을 지원하여 본 연구의 R1 및 R 조건을 구현할 수 있다는 점이 선정의 실질

표 4. 분석 코퍼스 구성(IMG-01~IMG-12)
Table 4. Composition of the analysis corpus

IMG-01		IMG-02	
			
Gemini	T	Gemini	T
Macro	Geunjeongjeon	Detail	Muryangsujeon
IMG-03		IMG-04	
			
Gemini	R1	Gemini	R1
Macro	Hyangwonjeong	Detail	Geunjeongjeon
IMG-05		IMG-06	
			
Gemini	R	Gemini	R
Macro	Ganghakdang	Detail	Hyangwonjeong
IMG-07		IMG-08	
			
Midjourney	T	Midjourney	T
Macro	Muryangsujeon	Detail	Geunjeongjeon
IMG-09		IMG-10	
			
Midjourney	R1	Midjourney	R1
Macro	Geunjeongjeon	Detail	Hyangwonjeong
IMG-11		IMG-12	
			
Midjourney	R	Midjourney	R
Macro	Ganghakdang	Detail	Muryangsujeon

*Chinese characters are retained in their original form and are not relevant to the interpretation of this study

적 요건이었다. 연구 수행 시점을 기준으로 Stable Diffusion 및 DALL·E 등은 참조 이미지 활용 방식이 상이하여 본 연구가 설정한 조건을 동일하게 통제하기 어렵다는 연구 범위상의 이유로 제외하였다.

R1 및 R 조건에서는 원본 참조 이미지를 각 모델의 이미지 입력 인터페이스에 직접 업로드하여 입력하였으며, R 조건에서는 참조 이미지 입력에 추가하여 시점 지시구를 부가하였다. Midjourney는 v7을 사용하였고 생성 파라미터는 --ar 3:2, --raw, --v 7, --iw 2로 고정하였으며, 참조 이미지 가중치(--iw)를 2로 설정하여 참조 이미지 반영 조건을 일정하게 유지하였다. Gemini는 유료 구독 환경에서 이미지 생성 모델 Nano Banana 2를 사용하였다. 생성 및 이미지 선정은 모두 2026년 5월에 수행하였다. 두 모델은 성능 비교의 대상이 아니라 서로 다른 상용 생성 환경에서 전통 건축 이미지의 재현 특성을 탐색하기 위한 사례로 사용하였으므로, 결과는 통합하여 분석하였다.

각 조합에 대해 동일한 프롬프트로 반복 생성한 후, 프롬프트에 기술된 구조·장식 요소가 육안으로 확인 가능하며 후속 전문가 검증에서 공포, 결구, 단청 등 주요 요소를 판별할 수 있는 결과물 1장을 선정하였다. 선정 기준은 원본 이미지와의 유사도나 시각적 완성도가 아니라 프롬프트 지시의 반영 여부와 평가 가능성이었다. 다만 이 과정에서 연구자의 판단이 개입하였으므로, 본 연구의 결과는 생성 모델의 평균적 출력이 아니라 프롬프트 지시가 반영된 조건에서의 재현 특성으로 해석되어야 한다. 이와 같은 절차를 통해 건축물 유형(4) × 이미지 범위(2) × 생성 조건(3) × 생성 모델(2)의 각 조합에서 1장씩 총 48장의 분석 이미지를 구축하였다.

이후 도구(2) × 생성 조건(3) × 이미지 범위(2)의 12개 조합이 각 1장씩 반영되도록 표 4와 같이 총 12장의 대표 이미지를 선정하였다. 셀 내에서는 정량분석 결과를 참조하여 특정 수준의 사례에 치우치지 않도록 선정하였다. 그 결과 선정된 12장이 각 조합 내 네 개 사례에서 차지하는 구조적 유사성 순위는 1위 3건, 2위 2건, 3위 4건, 4위 3건으로 분포하였으며, 12장의 SSIM은 0.058에서 0.835에 이르는 범위를 포괄하였다.

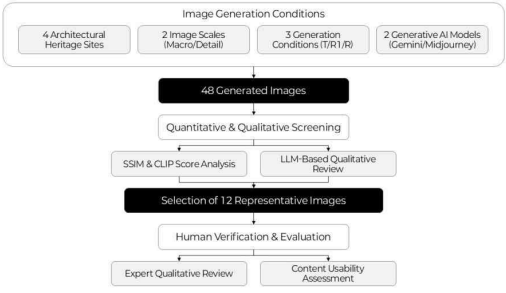


그림 3. 연구 절차
Fig. 3. Research workflow

전체 연구 절차는 (1) 전통건축 원본 이미지 수집, (2) 생

성형 AI를 활용한 48장 분석 이미지 구축, (3) SSIM·CLIPScore 기반 정량분석, (4) LLM 기반 예비 질적 평가, (5) 대표 사례 12장 선정, (6) 전문가 검증, (7) 콘텐츠 활용성 평가, (8) 결과 통합 및 용도-정밀도 매칭 프레임워크 도출의 순서로 진행하였으며 그림 3과 같다.

3-2 LLM-Human 다층 검증 프레임워크

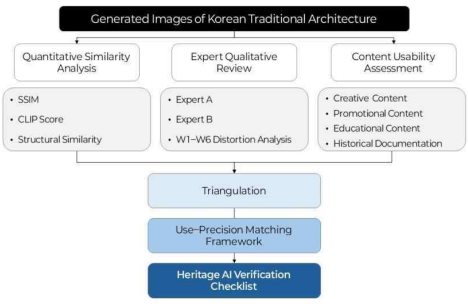


그림 4. LLM-Human 다층 검증 프레임워크
Fig. 4. LLM-Human multi-layer verification framework

본 연구에서는 생성형 AI 기반 전통 건축 이미지의 활용 가능성을 검증하기 위해 정량분석과 인간 평가를 결합한 다층 검증 체계를 적용하였다. 그림 4와 같이 이미지 생성, 정량 및 예비 질적 분석, 전문가 검증, 콘텐츠 활용성 평가, 삼각 검증의 다섯 단계로 연구를 수행하였다.

먼저 Claude 3.7 Sonnet을 활용하여 전통 건축의 주요 구조 요소를 반영한 프롬프트를 구성하고, 이를 Midjourney와 Gemini에 적용하여 총 48장의 분석 이미지를 구축하였다. 이후 SSIM과 CLIPScore를 활용하여 생성 이미지의 구조적 유사성과 의미적 정합성을 보조적으로 평가하였다. SSIM은 원본 및 생성 이미지를 512×512 pixels로 조정하고 PIL의 LANCZOS 방식으로 리샘플링한 후, RGB 3채널과 data range=255 조건에서 원본-생성 이미지를 1:1 비교하여 산출하였다. CLIPScore는 OpenAI CLIP ViT-B/32를 이용하여 생성 이미지와 프롬프트 간 이미지-텍스트 임베딩의 코사인 유사도로 산출하였다. SSIM은 R1에서는 구조 재현 비교에, T와 R에서는 조건 간 상대적 경향을 확인하는 보조 지표로 활용하였다.

그 다음, 멀티모달 LLM의 이미지 입력 기능을 활용하여 원본-생성 이미지 쌍을 비교하는 예비 질적 분석을 수행하였다. 예비 질적 분석은 표 3에 제시한 여섯 가지 왜곡 유형(W1~W6)을 평가 항목으로 하여 각 항목을 A(허용 가능), M(부분적 왜곡), H(심각한 왜곡)의 3단계로 판정하고 판정 근거를 자유 서술로 병기하는 방식으로 수행하였으며, 전문가 검증에 앞서 왜곡 유형의 발현 양상을 사전에 파악하기 위한 탐색적 절차로 활용하였다.

앞서 기술한 바와 같이 대표 이미지의 선정에는 설계 매트릭스와 정량분석 결과를 근거로 하였으며, 예비 질적 분석의 판정 결과는 선정 기준으로 사용하지 않았다. 실제로 선정된 12장의 종합 판정 분포(A 25.0%, M 58.3%, H 16.7%)는 48장 전체의 분포(A 27.1%, M 62.5%, H 10.4%)와 유사하여, 예비 질적 분석에서 상대적으로 양호하게 판정된 사례가 선택적으로 포함되었다고 보기 어려웠다.

한편 예비 질적 분석의 결과는 이후 전문가 검증과 항목에 따라 상이하게 나타났다. 문양 환각(W4)에서는 판정이 대체로 수렴하였으나, LLM이 전 항목을 허용 가능(A)으로 판정한 이미지에서 전문가는 활주의 누락을 지적하였고, 재료 표현을 허용 가능으로 판정한 이미지에서는 기단 재료의 오독을 지적하였다. 이는 LLM 기반 평가가 시각적으로 두드러지는 장식 요소의 이상은 포착하나 구조 부재의 누락과 재료의 물성 오독은 식별하지 못함을 보여준다. 따라서 본 연구에서 LLM은 프롬프트 구성, 정량 자료의 정리, 예비 질적 분석에 활용되었으며, 왜곡 판정과 활용 적합성에 대한 최종 판단은 모두 인간 평가자가 수행하였다.

표 5. 평가자 구성
Table 5. Evaluator profiles

Expert Qualitative Evaluation		
Evaluator	Area of Expertise	Role
Expert A	Certified National Heritage Repair Technician	Review of structural and restoration aspects of traditional architecture
Expert B	Traditional Architecture Specialist	Review of architectural form, decorative elements, and visual representation

Content Usability Evaluation		
Category	Sub-field	n
Game	Concept art, planning and development	3
Comics/Animation	Webtoon, animation	3
Design	Graphic and marketing design	2
Content Production & Planning	Service planning, content production	4
Research/ Education	HCI research, art & technology/ content education	4
Art	Media art, oriental painting, folk painting	4
Total		20

전통 건축 전문가 2인은 원본 이미지와 생성 이미지를 비교하며 여섯 가지 왜곡 유형(W1~W6)을 기준으로 12장의 이미지를 독립적으로 평가하고 자유 서술 의견을 제시하였다. 본 연구는 통계적 일반화보다 전통 건축의 구조적·양식적 특성에 대한 심층적 질적 판단을 확보하는 데 목적을 두었으며, 이에 실무 경력과 학술 전공이라는 상이한 관점을 대표하도록 평가자를 구성하였다(표 5). 두 평가자는 사전 협의나 합의 없이 독립적으로 평가하였으며, 종합 판정의 일치율은 83.3%(10/12), 여섯 개 왜곡 항목의 단순 일치율은 70.6%,

Cohen's kappa는 0.448(선형 가중 kappa 0.475), 왜곡 지적 여부를 이분화한 kappa는 0.593으로 나타났다. 불일치 사례는 합의값으로 조정하지 않고 두 평가자의 판단과 자유 서술을 모두 유지하여 상호 보완적으로 분석하였다. 자유 서술은 반복적으로 나타나는 결함 유형을 중심으로 범주화하고, 각 지적이 양측 공통인지 단독인지를 구분하여 표기하였다.

동시에 콘텐츠 및 디자인 분야 전공자와 실무 경험자 20명을 대상으로 활용성 평가를 실시하였으며 평가자 정보 역시 표 5에 제시하였다. 평가자에게는 원본-생성 이미지 쌍 12세트를 제시하고, 각 생성 이미지에 대해 창작 및 컨셉아트, 관광 및 문화유산 홍보, 교육 및 해설 자료, 학술 및 역사 고증의 네 가지 활용 목적별 적합성을 각각 5점 Likert 척도(1: 매우 부적합 ~ 5: 매우 적합)로 평가하도록 하였다. 이에 따라 이미지-평가자 단위의 응답은 240건(12장 × 20명)이며, 네 가지 활용 목적에 대한 개별 평정 값은 총 960개가 수집되었다. 동일 평가자가 네 가지 활용 목적을 모두 평가하는 반복측정(repeated measures) 구조이므로, 활용 목적 간 차이는 비모수 검정인 Friedman 검정으로 분석하고, Kendall's W를 효과 크기 지표로 산출하였다. 유의한 차이가 확인된 경우 인접 활용 단계 간 사후분석으로 Wilcoxon Signed-Rank Test를 실시하였다.

마지막으로 정량분석, 전문가 검증, 콘텐츠 활용성 평가 결과를 상호 비교하는 삼각 검증을 수행하여 평가 결과의 수렴 및 차이를 확인하고, 활용 목적별 요구 정밀도와 검수 수준을 체계화한 Heritage AI Verification Framework를 도출하였다.

Ⅳ. 연구 결과

4-1 정량분석 결과

본 연구에서는 SSIM(Structural Similarity Index Measure)을 생성 이미지와 원본 이미지의 절대적 일치도를 평가하기 위한 지표가 아니라, 생성 조건에 따른 구조적 재현 수준의 상대적 차이를 비교하기 위한 보조 지표로 활용하였다. 생성형 AI 이미지는 실사 이미지를 직접 복제하는 방식이 아니므로, SSIM 절댓값 자체보다는 조건 간 상대적 경향을 해석하는 데 의미를 두었다. 생성 조건 및 이미지 범위별 분석 결과는 표 6과 같다.

생성 조건별로는 참조 이미지를 활용한 동일 시점 생성(R1)이 가장 높은 구조적 유사성을 나타냈으며(SSIM=0.3491), 텍스트 프롬프트 기반 생성(T, SSIM=0.1514)과 참조 이미지를 활용한 시점 변환(R, SSIM=0.1840)은 서로 유사한 수준에 머물렀다. 반면 CLIPScore는 세 조건 모두 0.3453~0.3544 범위로 큰 차이를 보이지 않았다. 이는 생성형 AI가 전통 건축의 전반적인 형태와 분위기 등 의미적 특징은 비교적 안정적으로 이해하는 반면, 구조적 재현 수준은 참조 이미지가 동일 시점으로 제공되는지 여부에 따라 크게 달라짐을 보여준다.

표 6. 생성 조건 및 이미지 범위별 정량분석 결과
Table 6. Quantitative analysis results by generation condition and image scale

Category	Condition/ Scale	SSIM	CLIP	Interpretation
Generation Condition	T	0.1514	0.3544	Semantics understood, structural reproduction limited
	R1	0.3491	0.3460	Highest structural reproducibility
	R	0.1840	0.3453	Partial structural reproduction
Image Scale	Macro	0.2772	0.3581	Structure and semantics stable
	Detail	0.1792	0.3390	Reproduction strongly condition-dependent
Scale × Condition	Macro — T / R1 / R	0.2064 / 0.3806 / 0.2446	—	Condition effect within a narrower range (0.21–0.38)
	Detail — T / R1 / R	0.0965 / 0.3176 / 0.1234	—	Wider range across conditions (0.10–0.32)

이미지 범위별로는 전체(Macro) 이미지의 평균 SSIM이 0.2772인 반면 디테일(Detail) 이미지는 0.1792로, 세부 구조 요소의 재현 수준이 전체 형상에 비해 낮게 나타났다. 다만 이러한 격차는 생성 조건에 따라 상이하게 나타났다. 텍스트 기반 생성과 시점변환 조건에서는 디테일 이미지의 값이 전체 이미지의 절반 수준에 머물렀으나(T 0.2064 : 0.0965, R 0.2446 : 0.1234), 참조+ 동일 시점 조건에서는 그 격차가 크게 줄어들었다(0.3806 대 0.3176). 이에 따라 조건 간 차이의 폭 또한 전체 이미지(0.2064~0.3806)보다 디테일 이미지(0.0965~0.3176)에서 더 크게 나타났다. 즉 세부 구조의 재현 실패는 모든 조건에서 일관되게 나타나는 현상이 아니라, 참조 이미지가 동일 시점으로 제공되지 않는 조건에서 집중적으로 발생하였다. 다만 SSIM의 해석 범위에 관한 논의는 5-1절에서 다룬다.

4-2 전문가 검증 결과: 반복적 왜곡 양상

전통 건축 전문가 2인의 평가 결과를 종합한 결과, 생성형 AI는 한국 전통 건축의 전체적인 형태와 시각적 정체성은 비교적 효과적으로 재현하는 반면, 세부 구조 요소에서는 반복적인 왜곡을 보이는 것으로 나타났다. 특히 공포, 결구, 활주,

고주 등 주요 구조 요소가 단순화되거나 일부 부재가 생략되는 사례가 확인되었다. 일부 이미지에서는 마루 모서리를 지지하는 활주가 누락되거나(IMG-05), 추녀부의 서까래와 공포부가 뭉개져 표현되었으며(IMG-07), 고주가 절단되거나 공포 형태가 구조적으로 성립하지 않는 사례도 확인되었다(IMG-06, IMG-12). 또한 결구 부재가 생략되거나 장식적 형태로 대체되는 사례(IMG-08)가 나타났다.

장식 요소에서도 단청 문양의 변형(IMG-08, IMG-10), 창호 형태의 오류(IMG-06), 현판 문자의 변형 및 판독 불가 사례(IMG-08)가 확인되었다. 특히 IMG-08에서는 현판 문자의 오류와 함께 보가 생략되고 포의 문양으로 표현되는 현상이 관찰되어, 구조 부재와 장식 요소가 혼동되는 양상이 나타났다. 또한 기단 재료가 실제와 다르게 표현되는 재료 오독 사례(IMG-11)도 확인되었다. 이러한 결과는 생성형 AI가 건축물의 전체적인 형태는 비교적 안정적으로 재현하더라도, 세부 구조와 장식, 재료의 정확한 재현에는 한계가 있음을 보여준다.

전문가의 자유 서술 의견을 범주화한 결과, 사전에 설정한 구조 단순화(W3)와 문양 환각(W4)에서 각각 부재 누락(W3-1)과 현판 문자 왜곡(W4-1)이 세부 왜곡 유형으로 추가 도출되었다. 전문가 검증에서 확인된 반복적 왜곡 양상은 표 7과 같다.

표 7. 전문가 검증에서 확인된 반복적 왜곡 양상
Table 7. Recurring distortion patterns identified by traditional architecture experts

Distortion Type	Cases	Rater	Usage Considerations
Structural Simplification (W3)	IMG-05	Single	Expert review required for educational and documentation use
	IMG-07 IMG-08 IMG-12	Both	
Proportion Distortion (W2)	IMG-06	Single	Revision recommended for spatial interpretation materials
Element Omission (W3-1)	IMG-05	Both	Restricted use for structural explanation purposes
Pattern Hallucination (W4)	IMG-06	Single	Re-examination required for heritage interpretation content
	IMG-10 IMG-08	Both	
Signboard Distortion (W4-1)	IMG-08	Single	Revision recommended for promotional use
Material Misinterpretation (W5)	IMG-11	Single	Acceptable for creative purposes

Note. Cases lists images in which each distortion was most prominently observed. Both = identified by both evaluators; Single = identified by only one evaluator.

4-3 콘텐츠 활용성 평가 결과

표 8. 활용 목적별 적합성 기술통계(N=240)
Table 8. Descriptive statistics of usability by intended use (N=240)

Intended Use	Mean	SD	Note
① Creative & Concept Art	3.72	1.38	Highest suitability
② Tourism & Heritage Promotion	3.10	1.28	
③ Educational & Interpretive Materials	2.77	1.36	
④ Academic & Historical Documentation	2.42	1.30	Lowest suitability

콘텐츠 실무자 및 전공자 20인의 평가 결과, 표 8과 같이 적합성 점수는 창작 및 컨셉아트(M=3.72, SD=1.38), 관광 및 문화유산 홍보(M=3.10, SD=1.28), 교육 및 해설 자료(M=2.77, SD=1.36), 학술 및 역사 고증(M=2.42, SD=1.30) 순으로 나타났다.

활용 목적 간 차이는 통계적으로 유의하였으며(Friedman $\chi^2(3)=248.53$, $p<.001$), 인접 활용 단계 간 사후분석에서도 창작-홍보, 홍보-교육, 교육-고증 모두 유의한 차이를 보였다(Wilcoxon Signed-Rank Test, 모두 $p<.001$). Kendall's W는 0.345로 나타나, 활용 목적에 따른 적합성 차이가 중간 수준의 효과를 보이는 것으로 확인되었다. 또한 창작, 홍보, 교육, 고증 순의 적합성 순위는 평가 대상 12개 이미지에서 일관되게 나타났다.

생성 조건별 평균 적합성은 참조 이미지를 활용한 동일 시점 생성(R1)이 가장 높았으며(M=3.24), 시점변환(R)(M=2.90)과 텍스트 프롬프트 생성(T)(M=2.87)이 그 뒤를 이었다. 자유 서술 응답에서는 "틀린 그림 찾기 수준으로 비슷하다", "분위기 전달은 가능하다", "실제 구조는 다르다", "현판과 단청이 달라졌다"와 같은 의견이 반복적으로 제시되었다. 이러한 결과는 생성형 AI가 건축물의 전체적인 형태는 비교적 안정적으로 재현하더라도, 세부 구조와 장식, 재료의 정확한 재현에는 한계가 있음을 보여준다.

4-4 삼각검증 결과의 수렴과 평가 기준의 차이

정량분석, 전문가 검증, 콘텐츠 활용성 평가는 전반적으로 일관된 경향을 보였다. 참조 이미지를 활용한 동일 시점 생성(R1) 기반 이미지들은 전체와 디테일 범위 모두에서 높은 구조적 유사성과 긍정적인 전문가 평가를 받았으며, 콘텐츠 활용성 평가에서도 비교적 높은 적합성을 나타냈다. 반면 텍스트 기반 생성(T)과 시점변환 생성(R)의 디테일 이미지들은 구조적 유사성이 낮게 나타났으며, 전문가와 콘텐츠 집단 모두에서 상대적으로 낮은 평가를 받았다. 특히 그림 5와 같은 IMG-08과 IMG-10은 전문가 검증에서 가장 낮은 종합 평가를 받았고, 콘텐츠 활용성 평가에서도 낮은 점수(IMG-08 평

균 2.49)를 기록하여 세 자료원 간 평가 결과가 가장 일관되게 수렴한 사례로 확인되었다.



*Chinese characters are retained in their original form and are not relevant to the interpretation of this study

그림 5. 저평가 대표 사례(IMG-08, IMG-10)

Fig. 5. Representative low-rated examples (IMG-08 and IMG-10)

다만 세 자료원은 이미지의 상대적 우열에 대해서는 대체로 일치된 경향을 보였으나, 평가 기준과 엄격성에서는 뚜렷한 차이를 나타냈다. 전통 건축 전문가들은 공포, 결구, 현판 등의 구조적 정확성과 역사적 진정성을 중시하여 일부 디테일 이미지에 대해 창작 목적까지 포함한 부적합 판정을 내린 반면, 콘텐츠 실무자와 전공자들은 동일한 이미지라도 전체적인 분위기와 시각적 완성도가 유지되는 경우 창작 및 컨셉아트 용도로는 활용 가능하다고 평가하였다. 이러한 차이는 생성형 AI 문화유산 이미지의 활용 가능성이 단일한 품질 기준만으로 결정되기보다, 평가 주체의 전문 영역과 활용 목적에 따라 달라질 수 있음을 보여준다. 따라서 생성형 AI 문화유산 이미지는 활용 목적에 따라 요구되는 검증 수준을 차등적으로 적용할 필요가 있으며, 이는 본 연구에서 제안하는 Heritage AI Verification Framework의 설계 근거로 활용되었다.

V. 논 의

5-1 생성형 AI 전통건축 이미지의 활용 특성

본 연구 결과, 생성형 AI는 전통 건축의 전체적 외형과 시각적 정체성은 비교적 안정적으로 재현하는 반면, 공포 구조, 결구 방식, 현판 문자, 단청 문양 등 세부 구조·장식에서는 반복적인 왜곡을 보였다. 이러한 재현 수준은 생성 조건에 따라 차이를 보였으며, 참조 이미지와 동일 시점을 활용한 조건에서 상대적으로 높은 건축적 완성도와 활용 적합성이 나타난 반면, 텍스트 기반 및 시점변환 조건에서는 일부 구조적·장식적 왜곡이 두드러졌다. 이는 생성형 AI가 전체 외형은 비교적 안정적으로 재구성하지만 정밀한 구조 규칙이 요구되는 영역에서는 시각적 인상에 의존한 생성을 수행함을 시사한다.

또한 전문가와 콘텐츠 집단은 동일 이미지에 대해서도 서로 다른 활용 기준을 적용하였다. 전문가는 구조적 정확성과 역사적 진정성을 중심으로 평가한 반면, 콘텐츠 실무자는 활용 목적에 따라 시각적 완성도와 분위기 전달이 확보된 경우 상대적으로 넓은 활용 가능성을 인정하였다. 특히 콘텐츠 집단의 적합성은 창작에서 고증으로 갈수록 감소한 반면, 전문가는 활용 목적과 관계없이 양식적·구조적 타당성을 중심으로 판단하였다. 따라서 용도–정밀도 매칭은 전문가의 판단을 기술한 것이 아니라, 전문 지식이 없는 콘텐츠 실무자가 활용 목적에 따른 채택 여부를 판단하도록 지원하는 절차로 이해할 수 있다.

정량 지표의 해석에는 한계가 있다. SSIM은 시점과 구도가 다른 원본–생성 이미지에서 구조적 왜곡과 구도 차이를 분리하기 어렵다. 실제로 시점변환 조건의 SSIM은 텍스트 기반 조건과 유사한 수준에 머물렀으며(전체 0.2446 대 0.2064, 디테일 0.1234 대 0.0965), 이는 시점 변화에 따른 구도 차이가 참조 이미지의 이점을 상쇄했을 가능성을 시사한다. 따라서 SSIM은 시점이 통제된 R1 조건에서는 구조 재현의 비교 지표로, T 및 R 조건에서는 보조 지표로 제한하여 해석하였다.

또한 R 조건은 기저 시점 서술과 추가 시점 지시가 중첩되어 있어 순수한 시점 변환 효과로 해석하기 어렵다. 다만 전문가가 지적한 활주·고주 누락, 공포의 구조적 불성립, 기단 재료 오독 등은 시점 지시 자체와 직접 관련되지 않았으며, 유사한 왜곡이 다른 조건에서도 확인되었다. 후속 연구에서는 기저 시점 서술을 제거하고 단일 시점 지시만을 적용하여 시점 변인을 독립적으로 통제할 필요가 있다.

5-2 Heritage AI Verification Framework

이상의 분석 결과를 종합하여 본 연구는 활용 목적에 따른 검수 체계인 Heritage AI Verification Framework를 도출하였다. 표 9는 프레임워크를 구성하는 각 검수 항목의 실증적 근거와 창작(Cr), 홍보(Pr), 교육(Ed), 고증·기록(Doc) 목적에 따른 검수 수준을 제시한다. 기존 문화유산 AI 관련 연구 및 윤리 프레임워크가 원칙과 거버넌스 차원을 다룬 것과 달리[4],[5],[10], 본 프레임워크는 개별 생성 이미지의 채택 가능성을 판단하는 실무적 검수 단위에 초점을 둔다. 검수 항목을 활용 목적에 따라 차등 적용함으로써, AI를 어떻게 도입할 것인가보다 생성된 이미지를 어떤 목적으로 사용할 수 있는가에 답하는 것을 목표로 하며, 원칙 수준의 규범과 현장의 채택 판단 사이를 연결하는 실행 단위의 도구로 위치한다.

표 9. 검수 항목별 도출 근거 및 활용 단계 매트릭스
Table 9. Empirical basis and purpose-specific verification matrix for the heritage AI verification checklist

Verification Item	Empirical Basis (Summary of Findings)	Cr	Pr	Ed	Doc
Building Identifiability	R1 macro images rated suitable in all categories (IMG-03, 04) → identity retained at overall-form level	●	●	○	○
Spatial Composition	Macro SSIM 0.2772 vs. Detail 0.1792; detail reproduction rises to 0.3176 under same-view reference → additional verification required for viewpoint-change cases	△	●	●	●
Structural Elements	“Merged masses” and “structurally impossible” forms were noted in IMG-07 and IMG-12, while omission of a major structural member (IMG-05) was also identified → recurrent structural distortion	–	△	●	●
Joinery Method	"Beam omitted and rendered as bracket pattern" (IMG-08) → structure replaced by decoration	–	–	○	●
Signboard Text	Signboard text was altered or rendered illegible in IMG-08.	–	○	●	●
Dancheong Pattern	Dancheong variation was identified in IMG-08 and IMG-10; pattern changes and omissions were repeatedly observed.	–	△	○	●
Material Representation	"Platform rendered as cement" (IMG-11) → frequent material misreading in detail images	–	–	△	●
Expert Approval	Images rated wholly unsuitable by experts were rated usable by content evaluators → judgment gap	–	–	○	●

Note. Cr = Creative use; Pr = Promotional use; Ed = Educational use; Doc = Documentation use.

● = mandatory verification;
○ = recommended verification;
△ = conditional verification depending on the intended use;
– = not required.

검수 항목의 도출 논리는 다음과 같다. 첫째, 참조 이미지와 동일한 시점에서 생성된 거시 이미지(IMG-03, IMG-04)에서 전체 형태와 건축물의 정체성이 비교적 유지된 결과를 바탕으로 건물 식별성을 창작 목적의 최소 검수 항목으로 설정하였다. 또한 시점 변화에 따라 공간 배치와 비례가 달라지는 점을 고려하여 공간 구성을 홍보 단계부터 검토하도록 하였다. 둘째, 공포 구조의 단순화 및 구조적 불성립(IMG-07, IMG-12), 결구 부재의 누락 및 대체(IMG-08), 활주·고주 등 주요 부재의 왜곡(IMG-05, IMG-06)이 확인됨에 따라 구조 부재와 결구 방식을 교육 및 고증·기록 목적의 주요 검수 항목으로 설정하였다. 셋째, 현판 문자(IMG-08), 단청 문양(IMG-08, IMG-10), 재료 표현(IMG-11)의 오류를 바탕으로 문자·문양·재료 항목을 활용 목적별 검수 체계에 포함하였다. 마지막으로 전문가와 콘텐츠 실무자·전공자 집단 간 평가 차이를 고려하여, 고증·기록 목적에서는 전문가 검토를 필수 절차로 설정하였다.

본 프레임워크는 용도-정밀도 매칭(Use-Precision Matching) 원리에 기반한다. 활용 목적이 창작에서 홍보, 교육, 고증·기록으로 이동할수록 요구되는 재현 정밀도는 높아지고 허용 가능한 왜곡의 범위는 감소한다. 다만 참조+ 동일 시점 조건에서 관찰된 높은 구조적 유사성은 재현 능력의 향상이라기보다 참조 이미지의 복제에 가까울 수 있으므로, 고증·기록 목적에서는 원본 자료를 직접 활용하는 편이 합리적일 수 있다. 생성형 AI의 실질적 이점은 원본이 존재하지 않는 시점이나 장면을 구성하는 경우에 나타나며, 바로 그 조건에서 왜곡이 가장 크게 발생한다는 점이 본 연구의 핵심적 시사점이다.

Ⅵ. 결 론

본 연구는 생성형 AI가 생성한 한국 전통 건축 이미지를 대상으로 정량분석, 전문가 검증, 콘텐츠 활용성 평가를 수행하고, 활용 목적에 따른 요구 정밀도를 탐색적으로 검토하였다.

첫째, 생성형 AI는 전통 건축의 전체적 인상과 시각적 정체성은 비교적 효과적으로 재현하였으나, 본 연구의 사례에서는 공포 구조, 결구 방식, 현판 문자, 단청 문양 등 세부 구조 요소에서 반복적인 왜곡이 확인되었다(RQ1). 둘째, 디테일 이미지의 구조 재현 수준은 생성 조건에 따라 뚜렷한 차이를 보였으며, 참조 이미지가 동일 시점으로 제공되지 않는 조건에서 상대적으로 낮은 재현 수준이 나타났다(RQ2). 셋째, 활용 목적별 적합성은 창작, 홍보, 교육, 고증 순으로 낮아졌으며 이러한 순위는 대표 이미지 12개 이미지에서 일관되게 나타난 반면, 전문가는 활용 목적과 무관하게 단일한 타당성 기준을 적용하는 경향을 보였다(RQ3). 넷째, 이를 바탕으로 생성형 AI 문화유산 이미지를 단일 품질 기준이 아닌 활용 목적 기반으로 평가하는 Heritage AI Verification Framework를 제안하였다(RQ4). 다만 이상의 결과는 네 개 건축물 사례에서 관찰된 양상으로, 한국 전통 건축 전반이나 모든 생성형 AI 모델에 대한 일반화로 확대 해석되기 어렵다.

본 연구는 생성형 AI 문화유산 이미지의 평가 관점을 생성

성능 중심에서 활용 목적 기반으로 확장하고, 전문가와 비전문가 간 판단 괴리를 확인함으로써 검증 없이 유통되는 AI 이미지의 정보 왜곡 위험을 제시하였으며, 활용 목적별 필수 검수 항목을 통해 콘텐츠 디자이너와 문화유산 실무자가 AI 생성 이미지의 채택 여부와 검수 수준을 판단할 수 있는 실무적 기준을 제공하였다.

그러나 본 연구는 네 개의 전통 건축 사례와 두 명의 전문가, 특정 시점의 생성형 AI 모델을 대상으로 한 탐색적 연구라는 한계를 가진다. 또한 시점변환 조건에서 시점 지시의 중첩이 발생하여 시점 변인이 완전히 통제되지 않았으므로, 해당 조건의 결과는 제한적으로 해석되어야 한다. 구체적 성능 수치는 모델 업데이트에 따라 달라질 수 있으나, 활용 목적에 따라 검증 수준을 차등 적용하는 프레임워크의 구조적 원리는 유효할 것으로 판단된다. 향후 연구에서는 기저 프롬프트의 시점 서술을 제거한 단일 시점 지시 설계, 다양한 문화유산 유형과 이해관계자를 포함한 검증, 허용 왜곡 수준의 정량화 및 자동 검증 시스템 구축으로 확장할 필요가 있다.

부 록

표 10. 베이스 프롬프트

Table 10. Base prompts

Geunjeongjeon — Macro
Geunjeongjeon (勤政殿), the main throne hall of Gyeongbokgung Palace, Seoul, South Korea, viewed from the front along the central axis. A massive two-story wooden structure with a double-layered Paljakjibeung(hip-and-gable roof) — the upper roof broader and the lower roof projecting forward, both with strongly upswept eave corners and gray clay tile surfaces with white-plastered ridge crests. Beneath the eaves, dense rows of Dapo-style bracket sets (multi-cluster bracket system) are visible in polychrome Dancheong painting of green, red, blue, and gold. The upper story shows a horizontal signboard with gold Chinese characters (勤政殿) on a dark background. The facade displays tall reddish-brown wooden columns between which green-painted geometric lattice doors (Kkotsal-mun) fill each bay on both levels, with a continuous green horizontal band between stories. The building stands on a two-tier elevated stone platform (Woldae) with carved stone balustrades, corner animal sculptures, and a central stone stairway (Dapdori) with carved dragon-pattern relief panels. The overall proportion is powerfully horizontal with a broad, symmetrical facade, the double roof creating a layered pyramidal silhouette. The stone courtyard (Jodang) extends forward with flat paving stones, and behind the hall, green mountain ridges and clouded sky are visible. Photorealistic, high detail,natural lighting.
Geunjeongjeon — Detail
Close-up of the upper facade and signboard area of Geunjeongjeon (勤政殿), Gyeongbokgung Palace, Seoul. The Dapo-style bracket sets (multi-cluster bracket system) are arranged in multiple tiers, stacked three to four layers deep, with interlocking bracket arms (Salmi) and bearing blocks (Soroban) painted in vivid Dancheong polychrome — greens, blues, reds, and golds — featuring Meoritcho petal-tip motifs on each rafter end and Bimeori-cho vine scrolls along the beam faces. The exposed rafter tails are arranged in a parallel pattern, each tip decorated with concentric colored bands. Between the two roof tiers, a row of small rectangular openings with wooden lattice grilles (Changbang, transom windows) allows light into the upper space, with rounded eave tiles (Makae) visible above. The central signboard displays three bold gold Chinese characters (勤政殿) on a black-lacquered ground, framed with a decorated border and flanked by red-painted bracket arms. The overall color palette is dominated by green beams, red-orange columns, and the multicolored floral medallion patterns on the horizontal members. All surfaces are painted wood with smooth, layered lacquer finish. Clear blue sky visible above the roofline. Photorealistic, high detail, natural lighting.

*Chinese characters are retained in their original form and are not relevant to the interpretation of this study

표 11. 이미지 생성 방법별 조건 설정

Table 11. Condition configuration by image generation method

Condition	Base Prompt	Reference Image	Viewpoint Directive
T (Text-to-Image)	Yes	—	—
R1 (Reference-Same View)	Yes	Yes	—
R (Reference-Novel View)	Yes	Yes	Appended (see Table 12)

표 12. 뷰포인트 지시 표현과 베이스 뷰포인트 프롬프트의 중첩

Table 12. Overlap between viewpoint deictic expressions and the base viewpoint prompt

Site / Scale	Viewpoint in Base Prompt	Appended Viewpoint Directive	Overlap
Geunjeongjeon (Macro)	viewed from the front along the central axis	viewed from the left at a 45-degree angle	Conflict
Muryangsujeon (Macro)	viewed from the front-left angle	viewed from the right at a 45-degree angle	Conflict
Ganghakdang (Macro)	viewed from the front	viewed from the rear side	Conflict
Ganghakdang (Detail)	Interior view looking outward from inside	looking outward through the lattice doors from inside	Redundant
Geunjeongjeon (Detail)	(none)	looking up at the dancheong and brackets from beneath the eaves	None
Muryangsujeon (Detail)	(none)	looking up diagonally at the entasis column	None
Hyangwonjeong (Macro)	(none)	viewed from across the pond	None
Hyangwonjeong (Macro)	(none)	looking up at the eaves and railing from below	None

감사의 글

이 논문은 2025년 대한민국 교육부와 한국연구재단의 지원을 받아 수행한 연구임(NRF-2025S1A5C3A04022639).

참고문헌

[1] I. Ibrahim, M. Abu Talib, A. Ammar, K. A. Tabet Aoul, and T. Abuimara, “Comparative and Experimental Analysis of Leading Text-to-Image Generative Artificial Intelligence Models for Regional Residential Architectural Designs,” *Results in Engineering*, Vol. 29, 108835, March 2026. https://doi.org/10.1016/j.rineng.2025.108835

[2] A. Odiah and S. D. Gosling, “Laying the Foundations for Using Generative AI Images in Architectural Research: Do Images Convey the Intended Spaces and Ambiances?,” *Architectural Intelligence*, Vol. 3, 35, October 2024. https://doi.org/10.1007/s44223-024-00076-x

[3] V. Paananen, J. Oppenlaender, and A. Visuri, “Using Text-to-Image Generation for Architectural Design Ideation,” *International Journal of Architectural Computing*, Vol. 22, No. 3, pp. 458-474, 2024. https://doi.org/10.1177/14780771231222783

[4] S. Münster, F. Maiwald, I. di Lenardo, J. Henriksson, A. Isaac, M. M. Graf, ... and J. Oomen, “Artificial Intelligence for Digital Heritage Innovation: Setting Up a R&D Agenda for Europe,” *Heritage*, Vol. 7, No. 2, pp. 794-816, February 2024. https://doi.org/10.3390/heritage7020038

[5] S. Pansoni, S. Tiribelli, M. Paolanti, F. Di Stefano, E. Frontoni, E. S. Malinverni, and B. Giovanola, “Artificial Intelligence and Cultural Heritage: Design and Assessment of an Ethical Framework,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLVIII-M-2-2023, pp. 1149-1155, June 2023. https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-1149-2023

[6] A. W. Sukkar, M. W. Fareed, M. W. Yahia, S. B. Abdalla, I. Ibrahim, and K. A. K. Senjab, “Analytical Evaluation of Midjourney Architectural Virtual Lab: Defining Major Current Limits in AI-Generated Representations of Islamic Architectural Heritage,” *Buildings*, Vol. 14, No. 3, 786, March 2024. https://doi.org/10.3390/buildings14030786

[7] C. Wei, J. Liu, J. Jia, D. Kong, M. Yuan, and S. Yan, “A Critical Artificial Intelligence-Generated Content Approach for the Reconstruction of Qing Palace Interiors: The Case of Juanqinzhai,” *npj Heritage Science*, Vol. 14, 124, February 2026. https://doi.org/10.1038/s40494-026-02390-1

[8] M. N. Akhtar, S. Iacono, G. V. Vercelli, H. Zulifiqar, and M. Suffian, “Generative AI for Digital Color Restoration in Archaeology: Reconstructing Lost Pigments of the Harappan Civilization,” *Digital Applications in Archaeology and Cultural Heritage*, Vol. 41, e00549, June 2026. https://doi.org/10.1016/j.daach.2026.e00549

[9] L. Tan and M. Luhrs, “Using Generative AI Midjourney to Enhance Divergent and Convergent Thinking in an Architect’s Creative Design Process,” *The Design Journal*, Vol. 27, No. 4, pp. 677-699, May 2024. https://doi.org/10.1080/14606925.2024.2353479

[10] G. Gattiglia, “Managing Artificial Intelligence in Archeology. An Overview,” *Journal of Cultural Heritage*, Vol. 71, pp. 225-233, January-February 2025. https://doi.org/10.1016/j.culher.2024.11.020

[11] J. W. Creswell and J. D. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 5th ed. Los Angeles, CA: SAGE Publications, 2018.

[12] J. W. Creswell and V. L. Plano Clark, *Designing and Conducting Mixed Methods Research*, 3rd ed. Los Angeles, CA: SAGE Publications, 2017.

[13] N. K. Denzin, “Triangulation 2.0,” *Journal of Mixed Methods Research*, Vol. 6, No. 2, pp. 80-88, April 2012. https://doi.org/10.1177/1558689812437186

[14] M. Q. Patton, *Qualitative Research & Evaluation Methods*, 4th ed. Thousand Oaks, CA: SAGE Publications, 2015.



김지은(Ji-Eun Kim)

2018년 : 경기대학교 애니메이션 영상학(예술학사)
2025년 : 중앙대학교 첨단영상대학원 예술공학(영상학석사)
2025년～현 재: 중앙대학교 첨단영상대학원 예술공학 박사과정
※관심분야 : Animation, Comics, Contents Design, Media Art 등



박진완(Jin Wan Park)

1995년 : 중앙대학교 컴퓨터 공학과 (공학사)
1998년 : Pratt CGIM ComputerMedia (MFA)
2003년～현 재: 중앙대학교 첨단영상대학원 예술공학 교수
※관심분야 : Art&Technology. Procedural Animation 등