2025년 추계학술발표대회 : 일반부문

LMM을 활용한 공공 건축물 인식 변화에 관한 기초 연구

A Basic Research on Changes in the Perception of Public Buildings utilising LMM

○박 가 은* 추 승 연**
Park, Ga-Eun Choo, Seung-Yeon

Abstract

This study explored the feasibility of evaluating public buildings using large multimodal models (LMMs). By analysing 10 cases of buildings with different evaluations in the past and present, five keywords were derived and LMM evaluation criteria were developed based on them. The experimental results showed that the LMM captured the changes in social perceptions and values reflected in architectural images to a large extent, but showed limitations in complex criteria such as 'innovation' and 'contextuality' that were inconsistent with existing evaluations. LMMs are valuable as an auxiliary tool for architectural evaluation, but they have limitations in understanding environmental context and need to be complemented by human qualitative evaluation. Future research is needed to design new prompts to overcome the context understanding limitations of LMMs.

키워드 : 공공 건축물, 평가, 대규모 멀티모달 모델, 건축 가치
Keywords : Public Architecture, Evaluation, Large Multimodal Models, Architectural Valuation

1. 서론

공공 건축물은 단순한 기능을 넘어 사회적·문화적 의미를 지닌다. 특정 건축물은 시대와 맥락에 따라 평가가 달라지며, 과거에 부정적으로 평가된 건축물이 오늘날에는 긍정적으로 재해석되기도 한다. 이러한 현상은 건축 평가가 문화적 배경, 시대적 가치관, 그리고 사회적 맥락에 따라 크게 좌우되는 복합성을 지니고 있음을 보여준다. 기존 건축 평가는 평가자의 주관성과 시대적 분위기에 따라 상이한 결과를 도출하며, 이는 건축물의 보존·활용·재개발 결정에 있어 혼란을 야기한다. 따라서 보다 일관된 평가 기준의 필요성과 다차원적 정보를 통합적으로 처리할 수 있는 새로운 평가 도구가 요구되고 있다. 최근 대규모 멀티모달 모델(Large Multimodal Model, 이하 LMM)의 발달은 텍스트와 이미지를 통합적으로 분석할 수 있어, 이를 활용하여 건축적 맥락을 분석과 건축 설계를 정량적 평가에 관한 연구가 활발히 진행되고 있다.

본 연구는 과거와 현재의 평가가 상이한 공공 건축물을 대상으로 LMM의 활용 가능성을 검증하고, 건축물의 시대적·사회적·환경적 맥락을 얼마나 이해하고 반영할 수 있는지 탐구하는 것을 목적으로 한다.

* 경북대 대학원 석사과정
** 경북대 건축학부 교수, 공학박사(Dr.-Ing.)
(Corresponding author : School of Architecture, Kyungpook University, choo@knu.ac.kr)
본 연구는 한국연구재단의 지원으로 수행되었음(NO. RS-2024-00349586).

2. 선행 연구 분석

건축 설계의 정성적 평가 부분을 정량화하기 위한 연구는 지속적으로 진행되고 있다. Lavdas&Salingaros(2022)는 건축의 객관적인 아름다움을 평가하기 위해 직관적인 반응을 점수화하는 Buras의 Beauty Scale를 바탕으로 실험자가 1차 평가를 시행하였으며, 이를 VAS(Visual Attention Scan) 분석에 결합하여 시각적 주목도와 건축의 조형적 특징을 분석하였다. Jiang B.(2025)는 'Living Structure' 이론 기반 직관적인 반응 기준인 Christopher Alexander의 Quality Without A Name(QWAN)에 기반하여 ChatGPT-4o를 사용하여 건축물이나 도시 공간의 아름다움을 평가하였다. Bingöl et al.(2025)은 ViT(Vision Transformer)와 SSA(Synergy Simulation Algorithm)을 활용하여 평가 기준의 유사도를 계산하여 건축 공모전 평가를 진행하였으며, 이미지에 스타일, 맥락, 디자인 원칙 기준에 대해 점수를 예측하는 다중 기준 점수 모델 평가도 시행하였다.

특히, 다중 기준 점수 모델 평가 방법이 이미지를 활용하여 스타일, 맥락, 디자인 원칙 등 복합적 기준을 종합적으로 평가할 수 있고, 건축물의 사회적·문화적 맥락을 보다 정확히 반영할 수 있어 본 연구에서 활용하기에 효과적일 것으로 판단되었다.

3. LMM 기반 공공 건축물 평가 및 분석

본 연구는 LMM이 시간에 따른 사회적 인식 변화와 건

축적 맥락의 복합성을 얼마나 이해할 수 있는지를 검증하는 것을 목적으로 한다. 이를 위해 CICA, RTF 등과 같은 여러 건축 비평지에 실린 사례 중 과거와 현재의 평가가 다른 10개의 건축물을 선정하였다. 이후 비평문에서 평가 항목별 언급의 빈도수를 확인하여 표1과 같이 분석하였다.

표1. 10개 건축물 평가 키워드 분석표

	혁신성	상징성	공공성	맥락성	기술적 진보	조형성	사용자 경험	기능적 제약	유지 관리 문제
1	2	2	1	1	1	1	2	1	0
2	1	1	2	1	0	1	1	2	1
3	2	2	1	2	1	2	0	0	0
4	2	2	1	1	1	1	0	0	0
5	1	2	0	1	0	2	0	1	0
6	2	2	1	1	1	1	0	0	0
7	1	1	2	1	0	1	0	0	1
8	1	1	2	1	0	1	0	0	1
9	2	2	1	2	1	1	0	0	0
10	2	2	1	2	1	1	0	0	0
합계	16	17	12	13	6	12	3	4	3

2점:반복적으로 언급, 1점:보조적이지만 반복적으로 언급, 0점:언급 없음

분석 결과, 빈도가 높은 순으로 상징성, 혁신성, 맥락성, 공공성, 조형성, 총 5가지 키워드를 도출하였다. 도출된 키워드를 바탕으로 기존 건축물을 평가하기 위한 5가지 프롬프트를 아래와 같이 정리하였다.

1. 상징성 - 도시나 사회에서 어떤 상징적 의미를 가지는가?
2. 혁신성 - 이 건축물은 기존 건축과 차별화되는 혁신적인 요소가 있는가?
3. 맥락성 - 주변 환경, 도시 맥락과 조화를 이루는가?
4. 공공성 - 대중에게 개방성과 접근성을 제공하는가?
5. 조형성 - 시각적, 형태적, 재료적 측면에서 실험적 디자인이 나타나는 정도는 어느 정도인가?

본 연구에서는 LMM 모델 중 가장 대중적인 ChatGPT-4o 모델을 사용하여 위에서 정리한 프롬프트를 건물 이미지와 함께 입력하였다. 1~5점의 척도로 프롬프트에 대한 점수를 표 2와 같이 도출하였다.

표2. GPT 평가 실험 결과

	혁신성		상징성		맥락성		공공성		조형성	
	과거	현재	과거	현재	과거	현재	과거	현재	과거	현재
DDP	5	4	4	5	2	3	3	4	5	4
파리 에펠탑	5	4	2	5	2	4	3	5	2	5
뉴욕 구겐하임 미술관	5	4	3	5	2	4	3	4	5	5
루브르 피라미드	5	4	2	5	2	4	4	5	4	5
시드니 오페라하우스	5	4	3	5	4	5	4	5	5	5

1점:미미함, 3점:중간, 5점: 확연함

GPT 실험 결과, 혁신성, 맥락성, 그리고 공공성에서는 1, 2점 차이로 미미한 결과가 나타났다. 하지만, 상징성과 조형성에서는 파리 에텔탑과 루브르 피라미드가 최대 3점까지 차이남을 알 수 있었다. 특히 가장 많은 분야에서 큰 차이를 보인 파리 에펠탑의 경우, 과거에는 도시경관 등을 고려하여 흉물이라는 평가가 있었을 만큼 상징적, 조형적 평가가 저조하였지만, 현재의 평가는 이와 다르게 나타났기 때문에 객관적인 판단기준이 부여된 GPT 평가에서는 과거와 현재의 평가가 차이가 나타났다.

각 건축물의 현재와 과거에 대한 기존 건축 평가와 GPT 평가를 표 3으로 정리하였다.

표3. GPT 평가와 기존 건축 평가 비교

	혁신성		상징성		맥락성		공공성		조형성	
	과거	현재	과거	현재	과거	현재	과거	현재	과거	현재
DDP	●	●	·	●	●	●	▲	●	▲	●
파리 에펠탑	●	●	●	●	▲	●	●	●	●	●
뉴욕 구겐하임 미술관	●	●	●	●	▲	▲	●	●	●	●
루브르 피라미드	▲	●	▲	●	▲	●	●	●	●	●
시드니 오페라하우스	●	●	▲	●	▲	●	●	●	●	●

●:완전 일치, ▲:부분 일치, ·:불일치

기존 비평과 GPT 평가를 비교하였을 때, 대부분 높은 일치도를 보였으나, 상징성과 맥락성은 부분적으로 불일치하는 평가가 많았다. 특히, DDP의 경우 상징성에서 과거의 주관적인 평가에 비해 객관적인 GPT 평가를 했을 경우 모든 분야에서 평가가 어긋남을 보였다.

4. 결론

본 연구는 과거와 현재의 평가가 상이한 건축물 사례를 중심으로 공공 건축물 평가 단계에서 LMM 활용 가능성을 검증하였다. 실험 결과, ChatGPT 평가는 건축물 이미지에 반영된 사회적 인식과 가치 변화를 상당 부분 포착했으나, 상징성·맥락성과 같이 환경적 판단이 중요한 기준에서는 기존 평가와 불일치하는 경우가 있었다.

결론적으로 본 연구는 ChatGPT와 같은 LMM이 건축물 평가에 대한 새로운 보조적 도구로서 유의미한 가능성을 가짐을 확인하였다. 향후 연구에서는 LMM 평가 한계를 극복하기 위한 환경 데이터와 맥락 정보를 보강하고, 보다 다양한 건축물을 대상으로 정성적 해석과 LMM의 정량적 분석을 연계하는 보완적 평가 체계가 모색되어야 한다.

참고문헌

1. Lavdas, A. A., & Salingaros, N. A. (2022). Architectural beauty: Developing a measurable and objective scale. Challenges, 13(2), 56.

2. Kaan Bingöl, Mustafa Koç, Selen Çiçek, Mehmet Sadık Aksu, Emre Öztürk, Gizem Mersin, Oben Akmaz, Lale Başarır (2025). Synthetic Interpretations: AI-Driven scoring framework for architectural design evaluation. eCAADe 2025

3. Jiang B. Beautimeter: Harnessing GPT for Assessing Architectural and Urban Beauty Based on the 15 Properties of Living Structure. AI. 2025; 6(4):74.

4. Peng Z.-R., Lu K.-F., Liu Y.H., and Zhai W. (2023), The pathway of urban planning AI: From planning support to plan-making, Journal of 480