



AI 이미지 생성기를 통한 건축 이미지 생성 실험 및 활용성 평가

Experiment and Evaluation of Architectural Image Generation through Artificial Intelligence-Based Text Image Generation Tool

이동호* · 고성학**

Dong Ho Lee* · Sung Hak Ko**

* Graduate Student, MSD-AAD, Univ. of Pennsylvania, Philadelphia, PA USA (postls95@gmail.com)

** Corresponding author, Assistant Professor, School of Architecture, Yeungnam Univ., South Korea (shko@yu.ac.kr)

ABSTRACT

Purpose: The accelerating speed of technological advancement means that AI technology cannot be isolated from our lives. And artists have already started to exchange a lot of information about using AI image generators. By comparison, there is a lack of research on AI image generators for architecture. Therefore, to highlight the possibility of an AI image generator based on a text-to-image model for application and utilization in actual architecture, the paper proceeded with the process of testing and evaluating. **Method:** Based on previous research, the paper outlined a new algorithmic methodology that is divided into a concept extraction phase and an architecturalization phase to generate creative concepts and architectural images. Each process of methodology was subjected to auxiliary experiments to ensure it worked effectively, and numerous image generation experiments were conducted to analyze and generate the result of methodology. **Result:** It was possible to derive and evaluate meaningful analysis results that can be used for other research and new methodology. And through a feedback loop process that transforms and repeats the entire process, it was possible to create images with creative colors and shapes that can be used in actual architecture. Finally, based on the limitations and possibilities of the current technology level, the paper predicted the future prospects of AI image generators in architecture and suggested the need for related follow-up studies.

KEYWORD

알고리즘
디자인 방법론
AI 이미지 생성기
프롬프트

Algorithm
Design Methodology
AI Image Generator
Prompt

ACCEPTANCE INFO

Received Sep. 6, 2023
Final revision received Oct. 12, 2023
Accepted Oct. 17, 2023

© 2023. KIEAE all rights reserved.

1. 서론

1.1. 연구의 배경 및 목적

인공지능(Artificial intelligence)은 최신의 기술이라는 일반적인 인식과 다르게 1940년대 후반 폰 노이만(John von Neumann)¹⁾에 의해 이미 그 가능성이 예측되었다. 특히 1960-70년대 MIT 인공지능 연구소를 설립한 마빈 민스키(Marvin Minsky), 인공지능(Artificial Intelligence)이라는 용어를 창안한 존 매카시(John McCarthy)와 같이 뛰어난 과학자들에 의해 학문적으로 정립되고 구체적인 방법론까지 도출되었으나, 당대의 물리적인 컴퓨팅 능력의 한계로 인해 유의미한 활용에 어려움을 겪었다. 그러나 최근 10년에 걸쳐 컴퓨터 연산 프로세싱 능력은 이전과 비교할 수 없는 가속성을 띠며 향상되었다. 자율주행 자동차, AI 생성 이미지의 예술 대회 수상, Chat GPT에서 촉발된 인공지능 경쟁 등 대중과 유리된 것으로 보였던 AI는 최첨단 기술의 선두 주자로서 과거의 상상을 현실화하며 점점 빠르게 우리의 일상 속으로 다가오고 있다.

AI가 근래 학술계를 넘어 일반 대중에게까지 폭발적으로 조명받는 이유는 컴퓨팅 기술의 발달로 이전까지의 프로그램과는 차별화된 뛰어난 성능을 보이는 것은 물론 AI에 대한 접근성이 크게 향상되었기 때문에 가능하다고 볼 수 있다. 자연어 처리 인터페이스와 함께 직관적인 UI로 제공되는 Chat GPT와 같은 생성형 AI 모델은 기존의 AI를 다루는 데 필요한 복잡한 사전 지식과 프로그래밍 언어에 대한 이해 없이도 누구나 사전 훈련된 뛰어난 성능의 인공지능에 쉽게 접근할 수 있도록 제공되고 있으며, Chat GPT의 급격한 가입자 수 증가량과 AI 시장을 선점하기 위한 구글과 마이크로소프트 간의 치열한 경쟁은 이에 대한 대중의 기대와 수요를 보여주는 수많은 지표 중 하나이다.

미드저니(Midjourney)를 포함한 AI 이미지 생성기의 경우, 2022년 디지털 예술 대회에서 인간의 창작물을 제치고 수상하는 것을 통해 성능에 대한 증명을 넘어 거대한 철학적 파동을 예술계에 촉발하였다. 강력한 컴퓨팅 자원을 기반으로 텍스트 세트와 대응하여 쌍을 이루는 이미지 세트의 대량 학습을 통해 유의미한 품질을 생산하는 AI 이미지 생성 기술은 Chat GPT와 마찬가지로 접근과 활용이 용이한 플랫폼의 형식으로, 1년도 지나지 않아 일반인은 물론 이를 적극적으로 활용하는 예술가도 등장할 정도로 사회, 예술계의 뜨거운 감자로서 수많은 논의를 재생산하고 있다[1, 2].

그러나 이와 관련된 건축계의 국내외 연구 사례는 손에 꼽을 정도로, 그 관심이 디자인 계통의 분야 중에서도 상대적으로 낮다고 할 수 있다. 그 이유로 1) 현재 접근 가능한 AI 이미지 생성기의 목적과 기술적 한계로 인해 2차원 이미지가 아닌 복잡한 3차원 건축 구조를

그러나 이와 관련된 건축계의 국내외 연구 사례는 손에 꼽을 정도로, 그 관심이 디자인 계통의 분야 중에서도 상대적으로 낮다고 할 수 있다. 그 이유로 1) 현재 접근 가능한 AI 이미지 생성기의 목적과 기술적 한계로 인해 2차원 이미지가 아닌 복잡한 3차원 건축 구조를

구축하는 것에 있어 명백한 한계를 가진다는 것, 2) 형성된 이미지는 반드시 학습을 위한 참조를 가지며 도시 문화적 복잡성과 건축적 창의성에 대한 논쟁이 있다는 것, 3) 최신의 도구로써 아직 실제 활용에 도달하기까지의 인지도, 그리고 성과를 확보하지 못하였기 때문으로 여겨진다.

그러나 AI 이미지 생성기의 성능과 수요는 이미 예술계의 사례를 통해 증명되고 있으며, 건축계에서도 급속도로 가속하는 현대 기술과 맞물려 예상보다 이른 시점에 현재의 CAD 및 BIM과 더불어 건축설계를 보조하는 주요 수단으로써의 요구가 가파르게 증가할 것으로 예상된다. 따라서 본 연구는 AI 이미지 생성 기술을 건축설계 과정에서 사용자와 결과물을 새로운 시각에서 이어주는 창의적 도구로 활용할 수 있도록 하는 방안을 모색하고자 한다. 특히 일반적으로 쉽게 접근할 수 있는 AI 이미지 생성기를 중심으로 건축적인 심미성을 충족할 수 있는 이미지 생성 전략을 도출하고, 나아가 실험의 과정과 결과를 평가하여 건축 디자인 과정에서의 AI 이미지 생성기의 활용 가능성을 크게 확대하고자 한다.

1.2. 연구의 방법 및 절차

본 연구는 Text-to-Image model의 일종인 스테이블 디퓨전 (Stable diffusion)²⁾ AI 이미지 생성기를 사용한다.³⁾ 먼저 AI 이미지 생성기와 관련된 선행 연구를 분석하여 창의적인 이미지를 생성하기 위한 초기 전략을 수립한다. 또한, 간략화된 프롬프트(Prompt)를 통해 학습 모델별로 전략을 실제 적용·실험하여 AI 이미지 생성기의 강점을 파악하고 이를 효과적으로 활용할 수 있는 구체적인 방법론을 구축한다. 이때 무작위의 2차원 이미지 생성에 특화된 AI 이미지 생성기가 방법론의 구성과 흐름 내에서 유의미한 관계를 끊임없이 유지할 수 있도록 img2img 기능과 프롬프트의 강도(Prompt strength)를 적극적으로 분석·활용한다. 최종적으로 가상의 목표를 설정하고 본 연구에서 구축한 방법론을 통해 실제 AI 이미지 생성 실험을 진행하여 현재의 AI 이미지 생성기가 목표를 달성하기 위한 건축 개념 및 입체적인 건축 이미지를 생성하는 것이 가능한지 실험하며, 전체 과정 및 결과에 대한 건축적인 유효성을 바탕으로 새로운 활용 가능성에 대해 평가한다.

2. 선행연구고찰 및 실험방법론 구축

2.1. 선행 연구 고찰

Table 1.에서 보는 바와 같이 현재 AI 이미지 생성기와 관련된 국내 연구는 아직 찾아보기 힘들며 그 내용도 현재 공개된 AI 이미지 생성기의 종류를 나열하는 간단한 소개 글에 머물고 있다. 한편 해외의 경우 관련 연구가 점차 증가하고 있으나 대다수가 짧은 견해(Opinion)를 포함한 건축 개념을 도출하기 위한 지원 도구로서의 가능성을 확인하는 정도에 그치고 있었다. 그러나 Jaruga-Rozdolska(2022)와 Radhakrishnan(2023)의 경우 AI 이미지 생성기에 대한 다양한 접근 방법을 제시하고 있었다. 따라서 창의적인 건축 이미지를 생성하기 위한 방법론 및 활용 전략을 구축하고자 하는 본연구의 기초 개념으로 활용할 수 있을 것으로 판단되어 다음과 같이 소개한다.

Table 1. Recent research on AI generator for architecture

Author (Year)	Title	Major Perspective
Jaruga-Rozdolska (2022)[3]	Artificial Intelligence as Part of Future Practices in the Architect's Work: Midjourney Generative Tool as Part of a Process of Creating an Architectural Form	3 approaches of prompt organizations to create AI architectural images
Ploennings & Berger (2022)[4]	AI Art in Architecture	Practical method and work flow based on quantitative look of 3 AI image generators
Dollens (2023)[5]	Stable Diffusion DALL-E 2, Midjourney and Metabolic Architectures	broaden possibilities of biological architecture through AI design
Radhakrishnan (2023)[6]	Is Midjourney-Ai the New Anti-Hero of Architectural Imagery & Creativity?	Educational approach of creative architectural thinking with AI
Wang & Cho (2022)[7]	Status and Potentials of Image-generator Artificial Intelligence System in Housing and Interior Design	Interior space generation through 3 AI image generators and comparison
Hwang & Cho (2022)[8]	A Study on The Artificial Intelligence Design Controlled by Design Language – Focus on Baby Schema design Language	Transforming linguistic concepts into a formative language through AI
Keyword : Chat GPT, Chat, GPT, Midjourney, Dalle2, Stable diffusion, Artificial intelligence, Image generator, Text to Image		

1) 건축 이미지 생성을 위한 프롬프트 구성에 관한 연구

Jaruga-Rozdolska(2022)는 미드저니를 기반으로 유효한 건축적 이미지를 생성하기 위한 프롬프트를 구성하는 3가지 접근 방법을 제시하였다. 첫 번째는 일반적인 Text-to-Image model의 사용 방식과 같이 목표로 하는 이미지를 직접 묘사하는 프롬프트를 사용하는 방식이었다. 해당 연구에서는 이것이 초기 개념을 직관적으로 구축하는 것에 도움을 줄 것으로 예측하였다. 두 번째는 바로크와 같은 디자인 양식과 관련한 프롬프트를 추가하여 구체적 지시가 없는 상황에서 형태를 구축하는 방식이었다. 실험의 결과물을 통해 Text-to-Image model이 건축 개념의 맥락까지 고려하여 이미지를 생성할 수 있는 잠재력이 있다고 판단하였다. 마지막은 건축 교육법에서 영감을 얻어 어떤 특정한 건축가의 스타일과 관련된 프롬프트를 작성하여 이미지를 생성하는 방식이었다. 비록 비전문적인 결과물이 나타났지만, 높은 수준의 그래픽을 가지는 색상 및 구성이 건축가의 지원 도구로써 충분히 활용될 가치를 가질 수 있음을 언급하였다. 결론적으로 현재 AI 이미지 생성기로 도출 가능한 결과물은 완전한 건축설계가 아닌 개념 스케치에 불과하나 건축가의 창의적 측면을 지원하고 초기 작업 단계를 가속하는 것에 큰 도움이 될 수 있음을 암시하였다[3].

2) AI가 건축의 창의적 사고에 미치는 영향에 관한 연구

Radhakrishnan(2023)은 인간의 창조과정과 미드저니 기반의 AI 이미지 생성과정을 구별하고 실제 AI의 결과물과 건축과 학생의 결과물을 비교하는 실험을 통해 건축 교육학적인 의의를 도출하였다. 연구에서는 실험에 앞서 AI 이미지 생성기에 대한 장점으로 우연적 창의, 빠른 속도, 다양한 변형을 언급하였고, 단점으로는 예술적 가치와 문화에 대한 이해 부족, 그리고 제한적 참조로 인한 한계를 지적하였다. 이어 추상적(예술적)인, 또는 공간적인 감정으로 구성된 프롬프트를 건축과 학생들과 AI 이미지 생성기에 제시하여 이미지를 생성하고 평가하는 실험을 진행하였고 다음과 같은 결론을 도출하였다. AI는 공간적인 감정의 프롬프트에 대하여 모호하고 부적절한 이미지를 생성하였으나 추상적인 키워드에서는 학생들보다 창의적인 평가를 받는 결과물을 생성해내었다. 이를 통해 AI가 직접 창의적인 건축 이미지를 생성하는 것에 명확한 한계가 있으나, 예술적 표현을 생성하는 것에는 유용한 창의적 점화 도구로 활용될 가치가 있다고 주장하였다. 이에 AI 이미지 생성기의 개념적인 건축 과정에서의 활용 가능성에 대해 탐색할 가치는 충분하며 이를 위해 Text-to-Image model이 건축적 창의성을 저해하지 않도록 안전하게 사용하는 방법에 관한 연구의 필요성을 주장하였다[6].

2.2. 실험방법론 구축

선행 연구를 참고하여 Text-to-Image model을 통해 창의적인 건축 이미지를 생성하기 위한 일련의 디자인 프로세스를 구축하고자 한다. 먼저 Radhakrishnan(2023)에서 도출된 AI 이미지 생성기의 창의적인 예술 표현에 대한 강점과 Jaruga-Rozdolska(2022)에서 제시된 건축적으로 구체화하기 위한 유효한 프롬프트 전략에 관한 아이디어를 최대한 독립적으로 활용하기 위해 프로세스의 함수 구조를 단계별로 나누어 개념을 도출하는 과정과 이를 건축적 이미지로 변환하는 두 가지 단계로 구분하여 진행한다. 특히 전 과정이 서로 다른 개별 생성 실험이 아닌 하나의 연속적인 프로세스로 기능할 수 있도록 스테이블 디퓨전의 img2img 기능을 적극적으로 분석하여 활용할 수 있도록 실험과정을 구축한다.

개념 도출 단계에서는 예술과 관련된 추상적인 키워드로 프롬프트를 구성하여 이미지를 생성한다. 따라서 직접적인 건축 관련 키워드 사용은 지양하나, 공간적인 촉매로써 감정4)을 보조 키워드로 사용하여 추상성에 대한 보강과 동시에 건축화 단계로 이어지는 판단 요소로 활용될 수 있도록 한다. 특히 Radhakrishnan(2023)에서 평가 집단을 모집하여 결과물을 분석한 것과 같이 추상적인 결과물의 평가 과정에는 실험자의 주관성을 최소화하기 위한 기준이 요구된다. 이에 미학 이론으로부터 미적 판단에 대한 틀을 마련하여 평가 및 분석을 진행한다. 동시에 최종 결과물을 통해 AI 이미지 생성기의 이해 능력과 생성을 지시한 사용자 의도 및 자연어 맥락에 대한 파악 능력을 판단한다.

건축화 단계에서는 img2img 기능을 통해 개념 도출 단계와 직접적인 연계성을 가지도록 설정한다. 특히 입력된 개념 이미지는 건축화 단계의 3가지 명령 수단인 직접적인 지시, 양식적인 설명, 건축적 예시를 포함한 프롬프트를 통해 의도하는 결과물로 변환된다. 이때 건축화 단계의 결과물은 직접적인 연계성을 갖는 미적 판단의 틀을

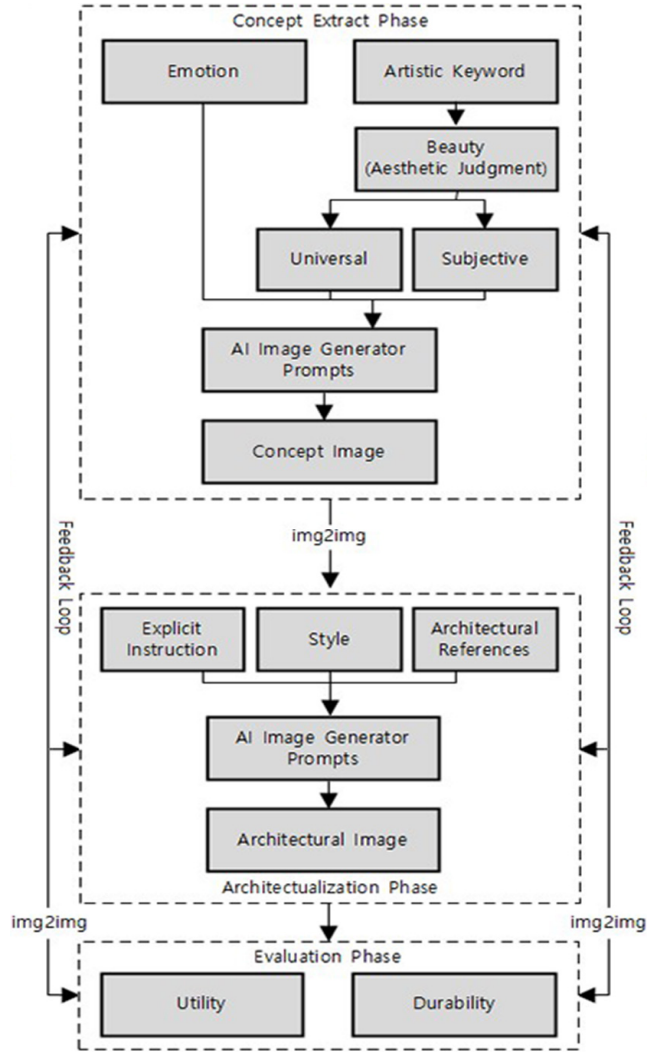


Fig. 1. Flowchart for image generation & evaluation

이미 통과하였기에 기초적인 건축 요소를 기반으로 평가되며, 의도한 목표를 달성하였는지도 중요한 평가 요소로 작용한다. 최종적으로 도출된 건축 이미지는 기존의 목표 또는 새로운 목표를 달성하기 위해 적절한 단계로 다시 환원되어 최선의 결과물을 얻기 위한 반복 과정을 거친다. 마지막으로 최종 결과물에 대한 분석 및 평가를 통해 현재 AI 이미지 생성기의 한계와 새로운 가능성을 파악한다.

2.3. 학습 모델 실험 및 평가

본 실험에 앞서 구축한 전략을 검증하고 결과물의 품질을 예측하기 위해 스테이블 디퓨전에서 사용할 수 있는 몇 가지 학습 모델을 추출하고 간략화된 방법론을 적용하여 실험에 적합한 모델을 선정하였다. 단 건축물만을 학습한 모델의 경우 학습 세트의 기호에 의한 편향성이 창의성을 저해하고 결과적으로 편견을 초래할 수 있으므로 실험에서 배제하였다. 개념 도출 단계와 건축화 단계에 해당하는 각 과정을 기초적인 프롬프트만을 통해 구현된 결과물 중 가장 생성 의도와 부합한다고 판단되는 이미지를 Fig. 2.에 나열하였다. 개념 도출 단계에서는 각 모델로부터 추상적으로 표현된 정원을 생성하도록 명령하였고, 건축화 단계에서는 img2img 기능을 통해 개념 도출 단계의 결과물을 작은 주택의 형태로 변환하도록 명령하였다.

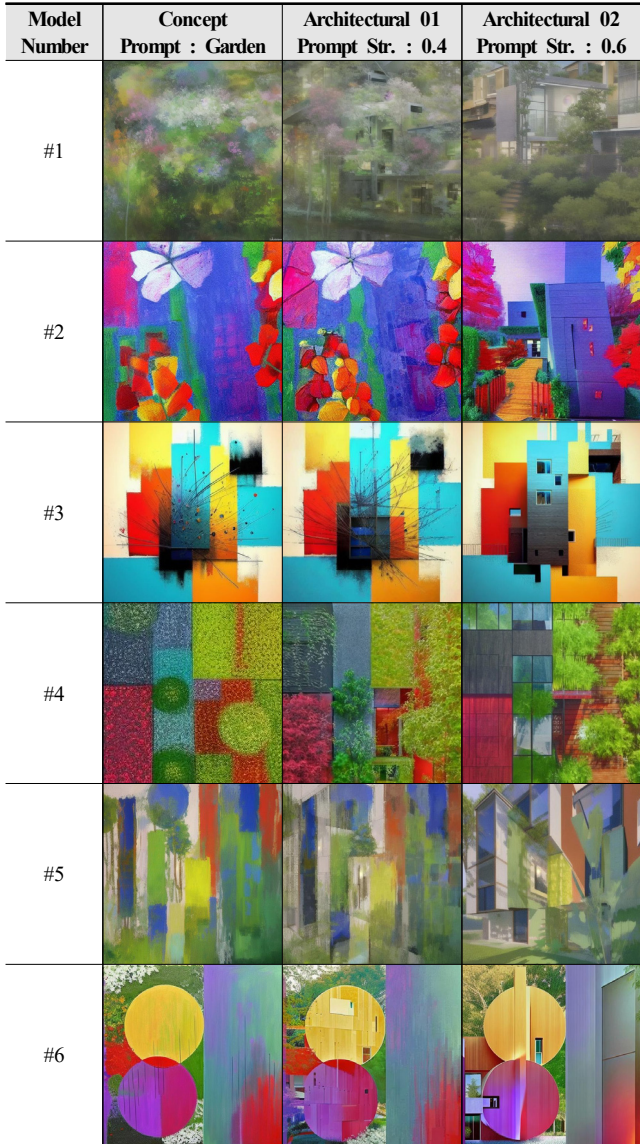


Fig. 2. Results of image models with stable diffusion

Fig. 2.에서 보는 바와 같이 학습 모델들은 프롬프트를 직관적으로 표현하는 이미지를 만드는 것에 강력한 성능을 발휘하였으나 이를 추상적인 결과물로 유도하기 위해서는 추가적인 프롬프트가 필요하였다. 프롬프트의 설정이 완료된 후에는 색감 및 품질이 준수한 다수의 결과물을 단시간에 생성할 수 있었으며 이를 건축적인 이미지로 변환하는 과정에서도 입력된 이미지의 색감을 잘 보존한 높은 품질의 이미지를 생성하였다. 다만 해당 실험에서는 명령 수단이 용도로만 한정되어 있었기에 건축물로서 아쉬운 결과물이 다수 생성되기도 하였다. 한편 실험 결과를 바탕으로 품질이 떨어져 실험에 사용되기에 부적절한 #2 와 #4 모델과 품질은 우수하나 편향 이미지 생성 비율이 높은 #1 모델을 배제하고 본 실험에서는 #3, #5, #6 모델만을 혼합하여 높은 품질의 결과를 얻을 수 있도록 의도하였다.

2.4. 이미지의 미적 판단에 대한 틀

일반적으로 어떤 대상의 아름다움(Beauty)에 관한 판단은 주관적인 경험이 배제된 대상 자체의 기하학적인 성질에 의해 결정될 수 있

다는 주장과 개인의 경험에 근거한 주관적인 만족으로 결정된다는 주장이 정반합의 관계를 맺으며 대립해왔다. AI 이미지 생성기의 예술적 창작에 주체성을 부여할 수 있는가에 관한 논의가 태동하는 현 시점에서 결과물이 자동 주체가 될 수 있음을 긍정하는 현대 미학 이론이 아닌 칸트의 판단력 비판(Critique of Judgement)에서 제안된 고전적인 미적 판단에 대한 틀을 본 논문에 활용하는 것이 실험의 결과를 객관적으로 평가하는데 적절하다고 판단하였다. 특히 해당 관점은 선행 규칙을 통해 AI 이미지 생성기의 결과물에 미적 가치를 부여하고, 예측할 수 없는 불가역적인 흐름을 가지는 AI 이미지 생성기의 특징과 부합된다고 판단하였다. 따라서 본 실험에서는 명백히 구분되는 고전적인 미적 판단의 형식만을 빌리되, 실례적인(Exemplary) 이미지의 재현 및 변형으로 귀결되는 예측 가능한 후자보다 추상적인 미적 보편성에 대한 실험에 집중하여 이상의 목적을 효율적으로 달성하고자 하였다[9, 10].

3. 건축 프로세스 실험 및 평가

3.1. 개념 도출 단계

개념 도출 단계는 인간의 창의적 사고절차를 활용하는 것은 물론, AI 이미지 생성기의 강점을 활용하여 Radhakrishnan(2023)에서 나타난 생성 이미지의 공간, 형태적인 한계를 극복하는 실질적인 절차의 기능을 갖는다. 특히 실험 주체인 프롬프트는 생성 규칙을 직접 지시할 수 있는 적극적인 창작 도구이자 불가역적인 진화 과정을 갖는 소통을 점화하는 촉매로써, 본 단계가 창의적인 건축을 위한 선행 단계이자 결과물의 미적 판단을 통한 가능성이 평가될 수 있도록 신중하게 구성되어야 한다. 이에 앞선 논의에 근거하여 비례, 조화 등 추상적인 미의 보편성과 관련된 키워드를 활용하여 이미지를 생성하는 실험을 진행하되, 대비, 대칭, 강조 등과 같은 구체적인 디자인 표현을 보조 키워드로 활용하여 직관적인 평가를 도울 수 있도록 하였다. 이와 더불어 감정의 키워드를 이들과 같은 강도로 활용하여 결과물이 앞서 언급한 미의 기준을 가진 이미지 생성은 물론 실험 주체의 의도와 소통하며 주관적인 맥락까지도 이해하고 표현할 수 있는지 확인하고자 하였다. Table 2.는 앞서 언급한 규칙을 바탕으로 키워드를 선정하고 이들을 결합하여 만든 기본적인 프롬프트의 예시이다.

Table 2. Prompt example for concept extract phase

(Beauty Component) + (Design component) + (Emotion)	
Selected keyword	Balance + Symmetry + Peaceful
Main prompts (Strength>Sub)	Balance, Symmetry
Sub prompts (Strength<Main)	Peaceful, Relaxed, Serene, Abstract art
Detail prompts	Artistic, Art, Realistic, 4k, High quality
Result	((balance)), ((symmetry)), ((peaceful)), ((relaxed)), ((serene)), (abstract art), artistic, art, realistic, 4k, high quality

특히 감정과 관련된 키워드는 본 연구의 또 다른 평가 지표로 각 집단과 독립된 변인으로 통제하여, 생성된 이미지의 주관적인 맥락 파악에 대한 평가가 가능하도록 고안하였다. 구체적으로 앞선 실험을 통해 선정된 3가지 학습 모델별로 제시된 키워드에 따라 8장의 이미지를 생성 및 분류하고 생성 이미지 세트들이 프롬프트를 통해 제시된 추상적인 미적 표현에 부합하는 적절한 이미지를 생성하였는지를 확인하는 과정을 거쳤다. 나아가 서로 다른 학습을 거친 각 모델 간의 차이를 극복하고 각 생성 이미지 세트별로 명확하게 구분되는 보편적인 의미를 도출하였는지 평가하는 동시에 서로 다르게 부여된 감정 표현이 주관적인 맥락에서 이해할 수 있도록 표현되었는지도 확인하였다. 이와 같은 실험은 다른 비례체계를 갖는 3차원 이미지에 대응하여도 동일하게 수행하였다.

생성된 이미지는 미적, 그리고 감정 표현에 관한 프롬프트에 따라 Fig. 3.과 같이 6개의 묶음으로 분류하였다. 각 이미지 세트는 묶음에 따라 유사한 관계성을 가지며 추상적인 표현에 대응하여 인간의 직관적인 인식에 부합하는 적절한 이미지를 생성한 것을 확인할 수 있었다. 특히 각 분류는 사용된 학습 모델과 관계없이 각 키워드에 대응한 공통점을 공유하고 있었는데, 균형과 대칭에 관한 결과물은 모두 어떠한 형태가 이미지의 중심으로부터 대칼코마니와 같이 방사되는 형태로 표현되고 있었으며 조화와 대비의 결과물들의 경우 뚜렷한 색 면과 이들의 분할이 주로 나타나고 있었다. 마찬가지로 감정 표현과 관련하여 차별함에 관한 키워드를 공유하는 이미지는 공통으로 원형의 푸른 파스텔 색조를 가지나, 강력함을 공유하는 경우 시인성이 높은 붉은색, 노란색, 검은색과 함께 직선적인 패턴이 주로 활용되는 등 각 분류에 따라 명백하게 구분 가능한 특징을 보여주었다. 이를 통해 AI 이미지 생성기가 추상적인 키워드의 맥락에 충분히 대응할 수 있으며, 과정을 반복하는 것을 통해 인간의 공통 감각(Sensus communis)과 유사한 방식으로 미의 보편성을 주관적인 맥락이 드러내는 결과물들을 통해 충분히 재현할 수 있음을 예측할 수 있었다.

반면에 같은 프롬프트로 다른 비례 체계 내의 3차원 이미지를 형성하도록 지시된 Fig. 4.의 경우 이미지의 품질은 우수하였으나 대부분 굉장히 부족한 공간감을 가지는 결과물을 도출하였다. 해당 결과물의 낮은 공간적 품질은 Radhakrishnan(2023)을 통해 어느 정도 예측할 수 있는 부분이 있었던 것은 사실이다. 그러나 키워드에 충실하게 대응하여 생성된 평면과 입체 사이에 걸쳐있는 결과물은 조립과 변형을 통해 창의적인 입면 디자인에 활용될 새로운 가능성을 충분히 내포하고 있다고 판단된다. 한편 정사각형이 아닌 비례체계는 이미지의 중심을 정의함에 있어 구체적으로 지시한 내용이 없어도 불구하고 수학적 중심에서 탈피한 우수한 비례의 결과물을 일부 생성하였다. 이는 건축 디자인의 다양한 비례체계에 대한 대응이 가능함을 암시할 뿐만이 아니라 앞선 공통 감각의 재현에 대한 논의를 강화해 줄 것으로 보인다.

다만 AI 이미지 생성기를 통해 연상되는 형이상학적인 현상은 인간의 지성적인 이해 과정과 동일시될 수 없으며, AI 이미지 생성기의 구조상 존 설(John Searle)이 제시한 중국어 방(The Chinese room)의 개념⁵⁾에 가까운 환각임을 유의해야 할 필요가 있다. 그럼에도 불구하고 창작 주체가 AI 이미지 생성기를 시스템이자 창의적

Emotion	Balance, Symmetry			Harmony, Contrast		
	#3	#5	#6	#3	#5	#6
Peaceful (Relaxed, Serene)						
Powerful (Important, Respected)						
Joyful (Excited, Energetic)						

Fig. 3. Results of prompts (512x512)

Emotion	Balance, Symmetry			Harmony, Contrast		
	#3	#5	#6	#3	#5	#6
Peaceful (Relaxed, Serene)						
Powerful (Important, Respected)						
Joyful (Excited, Energetic)						

Fig. 4. Results of prompts (3D Images)

인 창작 도구로써 의탁할 수 있으며, 합목적성의 표상에 대한 의도적인 착각을 통해 창작 주체의 미적 판단에 대한 가치를 부여할 수 있을 것으로 보인다. 다음 단계에서는 이러한 결과를 바탕으로 Fig. 3.에서 생성한 미적, 감정 표현과 관련한 특징들이 건축 이미지로 변환되었을 때 이를 충분히 보존한 창의적인 결과물을 생성하는 것이 가능한지 확인해 보았다.

3.2. 건축화 단계

건축화 단계에서는 개념 추출 단계의 결과물을 img2img 기능을 통해 실제 건축 디자인에서 활용할 수 있는 입체적인 이미지로 재생성하는 실험을 진행하였다. 특히 창의적인 결과물을 유도하기 위한 실험방법론의 의도에 따라 본 단계에서는 개념 이미지의 미적, 또는 감정적인 키워드와 관련한 특징들을 충분히 보존 및 활용할 수 있도록 유도되어야 했다. 이에 본격적인 실험에 앞서 감정별로 분류된 개념 이미지로 간략화된 선행 실험을 통해 본 실험의 품질을 높이기 위한 포인트를 도출하였다. 또한, 2장의 학습 모델 실험에서 확인한 img2img 기능의 프롬프트 강도를 선행 실험의 주요한 변인으로 사용하여 입력값에 따른 구체적인 관계성, 나아가 활용 가능성을 판단하였다.

Fig. 5.와 6.은 img2img 기능이 예배당(Chapel) 및 유치원(Kindergarten)과 같은 직접적인 프롬프트에 따라 추상적인 개념 이미지의 색과 형태를 충분히 보존하여 입체적인 건축 이미지를 생성한다는 사실을 보여주고 있다. 이는 감정 표현의 성질이 재생성된 건축 이미지에도 분명한 영향을 끼친다는 사실을 의미한다. 특히 예배당과 평온함(Peaceful), 유치원과 흥겨움(Joyful)과 같이 건축 용도와 유사한 맥락의 감정이 교차하는 경우, 다른 결과물에 비해 넓은 범위에서 픽진성이 증가하였다. 교차하지 않는 부분에서도 우수한 결과물을 일부 볼 수 있으나 활용할 수 있는 범위 및 생성 확률이 낮다는 측면에서 형태적 영향 이상으로 두 변인 간의 성질이 밀접하게 상호작용함을 알 수 있었다.

한편 언어적 명령 수행의 영향에 관여하는 프롬프트는 강도가 높아질수록 이와 같은 개념 이미지의 특징이 명암이나 재질 등으로 치환되며 열어지는 것을 확인할 수 있었는데, 이는 프롬프트 강도가 개념 이미지의 복잡도와 색감을 조정하기 위한 파라미터로 활용될 수 있음을 의미한다. 단, 프롬프트 강도는 실험 결과에서 보는 바와 같이 특정 이하의 값에서는 형태조차 구성하지 못하였으며, 반대로 특정 이상의 값에서는 사실상 개념 도출 단계가 없는 선행 연구와 같이 학습 세트의 영향만이 남은 특색 없는 결과물로 수렴하였다. 이에 실험 결과를 바탕으로 실험방법론의 목적 및 기존 선행 연구와의 차별성을 갖도록 최솟값과 최댓값의 유효한 범위를 한정하여 후속 실험을 진행하였다.

용도와 감정에 관한 상관관계 분석에 따라 예배당 및 유치원 각 용도의 합리적인 실험에 있어 상대적으로 더 넓은 프롬프트 강도 범위를 갖는 결과물을 Fig. 5.와 6.으로부터 추출하여 실험군으로 사용하였고, 이와 더불어 3가지 명령을 포함한 프롬프트를 함께 적용하여 전략의 실효성을 평가할 수 있도록 하였다. 구체적인 3가지 프롬프트 그룹은 다음과 같다. 먼저 용도, 규모, 재질, 형태 등을 명시적으로 지시하는 직접적인 프롬프트이다. 해당 프롬프트는 실험자의 의

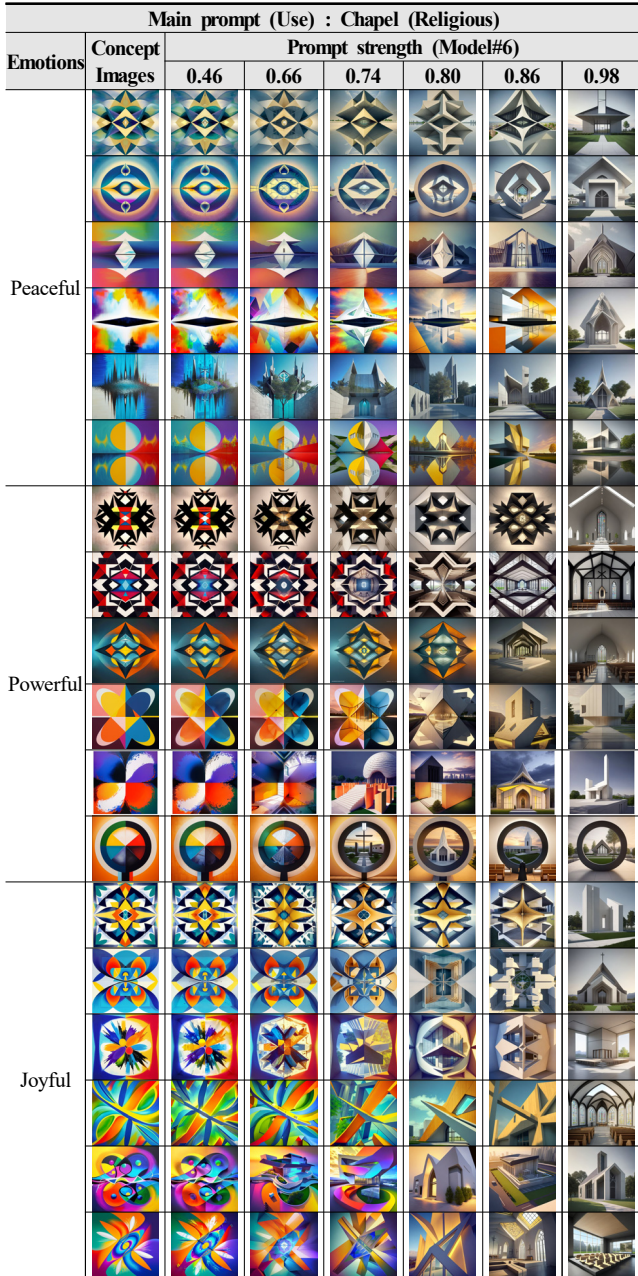


Fig. 5. Pilot experiment for architecturalization (Chapel)

도를 실현하기 위해 구체적으로 전달되는 명령이다. 두 번째는 Gothic, Modern, Futuristic, Biomorphic과 같이 이미지 전체의 시각적인 방향성을 결정하는 양식에 관한 프롬프트이다. 이것은 구체적 지시에 해당하는 첫 번째 명령보다 포괄적인 이미지의 분위기를 결정하는 것에 영향을 끼친다. 마지막으로 Le Corbusier, Zaha Hadid, Colosseum과 같이 건축가나 유명 건축물과 같이 건축적으로 참조 가능한 것들로 구성된 프롬프트이며 건축 이미지로서의 뒨진성을 강화한다. 해당 3가지 프롬프트는 기존 선행 연구와는 달리 결속된 하나의 프롬프트로써 실험에 사용되었다. 이를 통해 각 명령의 상호 영향성을 확인하는 것이 가능했으며, 실질적인 명령의 유효성 및 활용 가능성을 평가할 수 있게 되었다. Table 3.은 상기 내용을 기반으로 구성된 프롬프트의 예시이며 이를 바탕으로 서로 다른 목적을 가진 프롬프트를 두 세트 이상 구성하여 실험에 적용하였다.

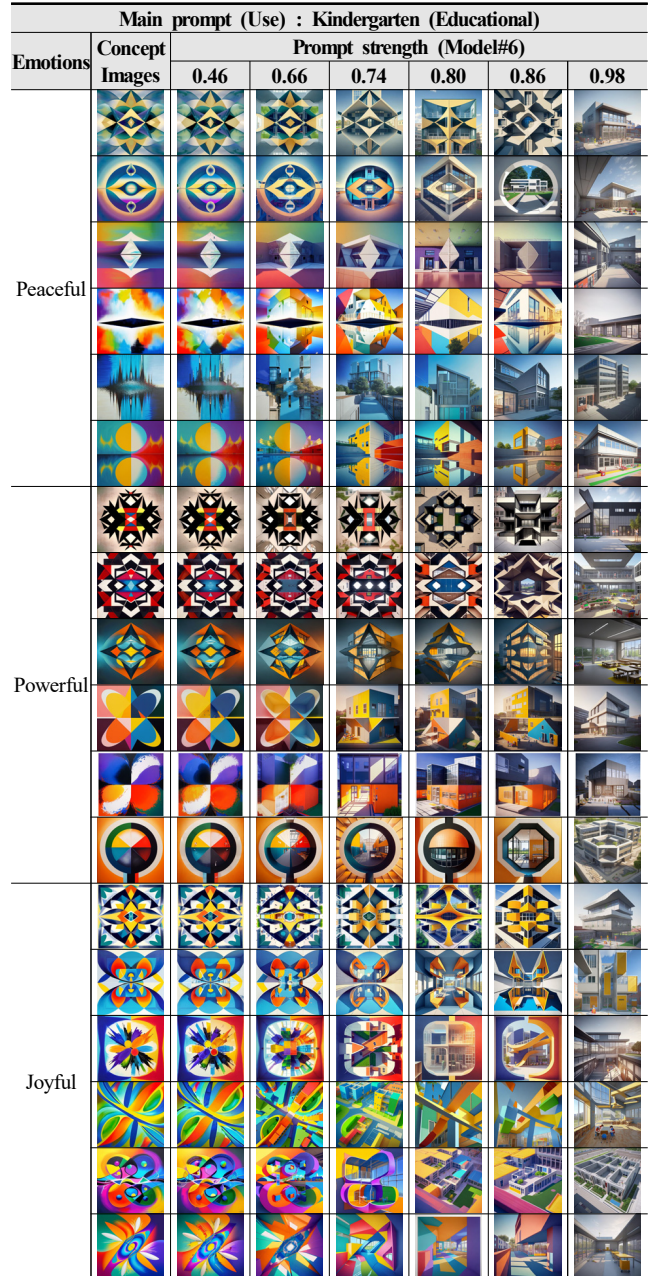


Fig. 6. Pilot experiment for architecturalization (Kindergarten)

Table 3. Prompt example for architecturalization phase

(Explicit instruction) + (Style) + (Architectural references)	
Main prompts	Chapel with terrace + Gothic + Zaha hadid
Detail prompts	Architecture, Architectural, Photograph, Photorealistic, Realistic, 8k, High quality
Result	((chapel with terrace)), ((gothic)), ((zaha hadid)), (architecture), (architectural), photograph, realistic, 8k, high quality

Fig. 7.과 8.은 3가지 명령에 각각 입력 가능한 서로 다른 2쌍의 키워드를 조합하여 4종류의 프롬프트로 압축한 후, 선행 실험에 따라 고정된 2가지 개념 이미지 및 용도 집합에 적용하여 비교분석이 가능하도록 정리된 실험의 결과물이다. 우선 Fig. 5.와 6.에서 각 개별 명령이 프롬프트로써 강력한 성능을 발휘한 것과 마찬가지로 Table

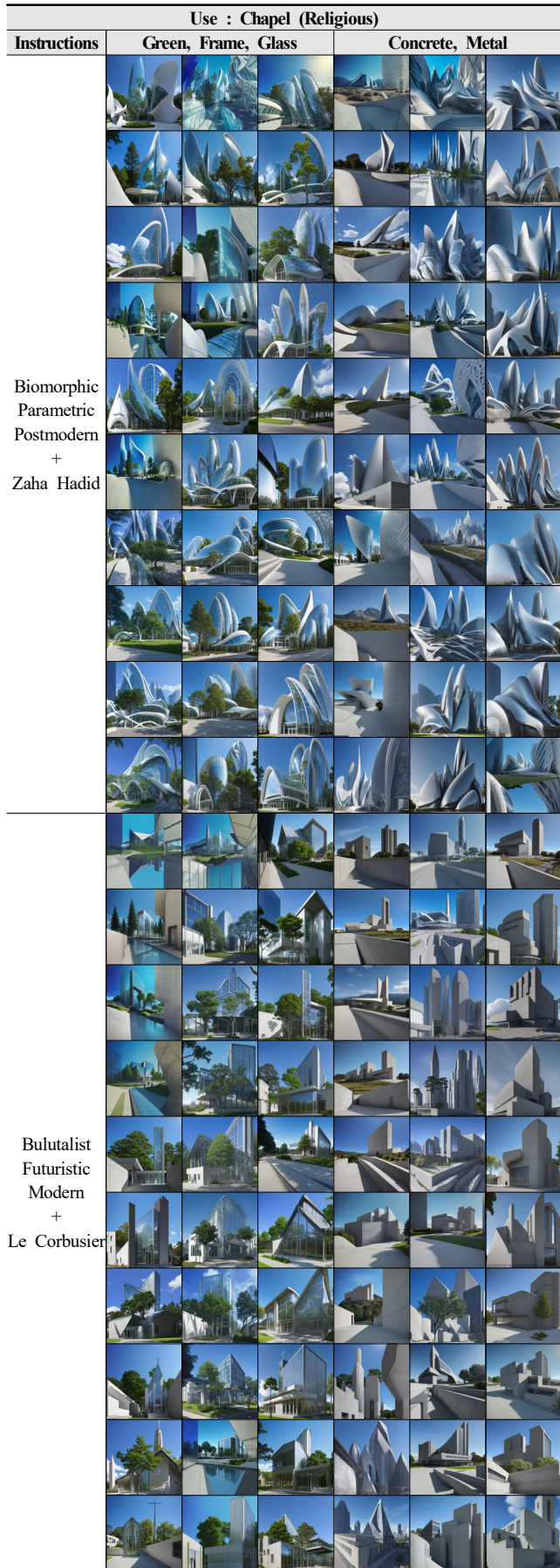


Fig. 7. Result of architecturalization phase (Chapel)

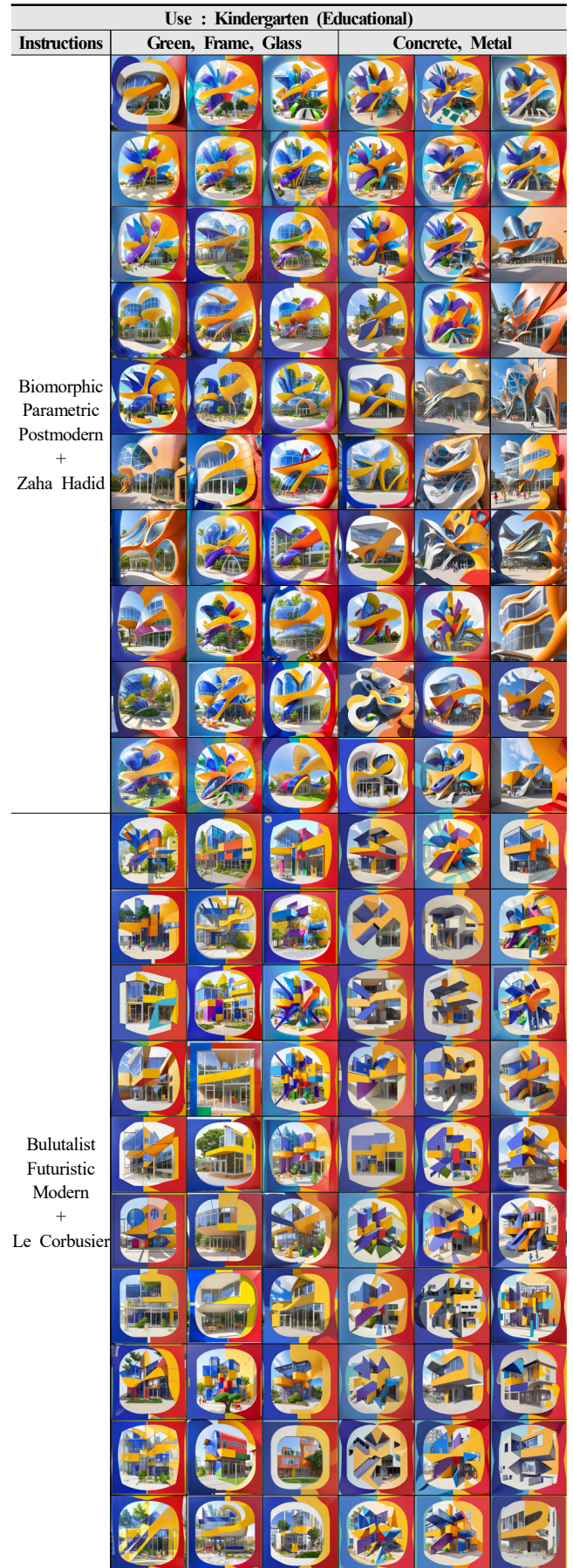


Fig. 8. Result of architecturalization phase (Kindergarten)

3.에 따라 조합된 프롬프트 역시 img2img 과정을 포함한 생성 결과물에 유효한 영향을 끼치고 있음을 확인할 수 있었다. 예를 들어 Fig. 7.에서 Green과 Glass의 키워드를 포함하여 생성된 이미지 집합은 프롬프트를 충실하게 반영하여 녹색의 자연적인 조경과 다량의 커튼월을 포함하고 있었다. 이들은 우측의 Concrete 및 Metal의 키워드를 포함한 이미지 집합과 명백하게 구분되며, 이는 Biomorphic, Zaha Hadid와 Bululalist, Le Corbusier의 결과물을 비교하는 것을 통해서도 확인할 수 있었다. 특히 의도와 근접한 이미지를 생성하기 위해 각 키워드 및 프롬프트의 강도를 세밀하게 조정하는 시행착오가 필요하였으나 가장 우려하였던 하나의 명령이 다른 하나의 영향력을 제거하는 경우는 발생하지 않았으며, 오히려 각 명령의 우선순위(강도)를 구체적으로 조절하여 실험자의 의도에 가까운 결과물의 생성 수단으로 활용할 수 있음을 확인하였다.

프롬프트 강도는 Fig. 6.의 이미지 분석에 따라 개념 이미지의 영향력을 통제하는 파라미터로도 활용하였다. Fig. 7.과 8.은 개념 이미지의 형태와 색이 각 키워드에 알맞게 변형 및 보존될 수 있도록 의도적으로 설정된 값에 의한 결과이며 우수한 품질의 결과를 통해 선행 연구에서 지적된 기존 생성 방식의 학습 모델로 인한 한계를 개념 이미지의 형태적 개성을 활용하여 극복하는 것이 가능함을 확인하였다. 다만 개성적인 형태는 이를 바탕으로 생성된 건축 이미지가 중력을 무시한 불안정한 형태의 이미지를 생성하도록 유도하기도 하였는데 이 또한 개별 강도 및 생성 개수의 조절을 통해 극복 가능함을 확인하였다.

최종 결과를 통해 전략 각 단계의 구축 의도는 충분히 목적을 달성한 것으로 판단된다. 이러한 결과는 본 프로세스를 AI 이미지 생성기의 창의성 및 편향성에 대한 약점을 극복하기 위한 하나의 방법론으로써 활용할 수 있음을 암시한다. 특히 건축 초기 단계는 물론, 창의적 도구로써 그 의도에 따라 자유롭게 결과물을 생성하는 것이 가능하기에 결과물 및 방법론의 변형을 포함하는 반복적인 개선 과정(Feedback Loop)을 통해 초기 단계 이후의 연속되는 건축의 각 단계에서도 적극적으로 활용될 수 있을 것으로 판단된다.

3.3. 피드백 루프 단계

실험을 통해 기존의 한계를 극복한 우수한 품질의 건축 이미지를 생성하는 것이 가능함을 확인하였다. 그러나 건축 디자인 과정에서 이상적인 최종 목표(Idea)의 완전한 실현은 사실상 존재하지 않으며, 이에 다가가기 위해서는 끊임없는 피드백 루프 과정을 반드시 수반하여야 한다. 무엇보다 피드백 루프 과정은 AI 이미지 생성기의 불확실성을 가장 효과적으로 통제할 수 있는 과정이며 결과물에 대한 개선을 넘어 전체 프로세스를 의미하는 방법론에 대한 새로운 전략 및 활용 가능성을 창출하는 것을 의미한다.

최종 결과물은 피드백 루프를 통해 끊임없이 변형 및 개선될 수 있다. 특히 이들은 실제 건축 과정에서의 각 활용 목적에 맞게 다양한 방법으로 수정되어 활용될 수 있으며 크게 언어적 개입과 시각적 개입으로 구분될 수 있다. 먼저 언어적 개입은 기존 이미지에서 개선하고자 하는 부분을 직접 프롬프트로 지시하여 변형하는 것을 의미한다. Fig. 9.와 같이 img2img 기능을 통해 전체 혹은 일부에 대한 변형을 유도할 수 있으며, 상대적으로 빠르게 많은 결과물을 도출할 수

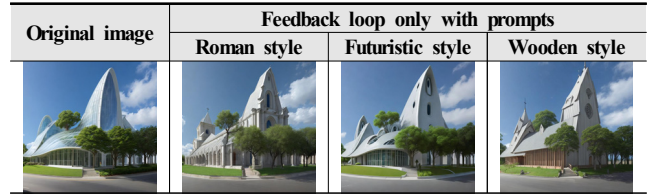


Fig. 9. Feedback loop example (Linguistic engagement)

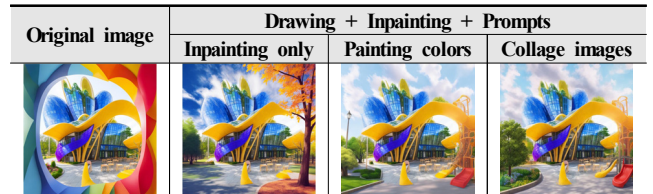


Fig. 10. Feedback loop example (Visual engagement)

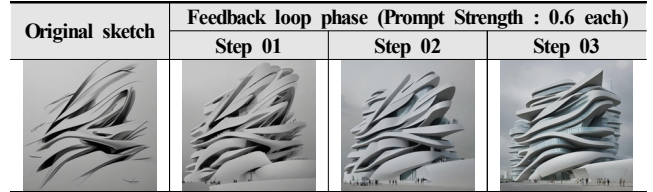


Fig. 11. Feedback loop example (Method modify)

있으나 AI의 불확실성에 의해 프롬프트의 구성 및 원하는 결과물을 정확하게 도출하는 것에 어려움을 겪을 수 있다.

시각적 개입은 Fig. 10.과 같이 직접 원형 이미지를 편집 및 img2img 기능을 통해 변형하는 것을 의미한다. 이러한 방식은 언어적 개입과 동시에 이루어져 편집된 이미지가 더욱 사용자의 의도에 가깝게 생성될 수 있도록 유도될 수 있다. 해당 방식은 언어적 개입만으로 진행되는 과정보다 구체적이고 정확한 결과물을 얻을 수 있다는 장점이 있으나 이미지를 수정하기 위해서 상당한 시간을 소모해야 하는 단점이 있다.

결과물이 아닌 방법론 자체를 활용하려는 각 건축 단계의 목적에 맞게 변형하여 새롭게 활용할 수도 있다. 예를 들어 개념 도출 단계에서 보편성에 대응하는 키워드가 아닌 또한 미적 판단의 틀에 포함되는 경험에 근거한 직관적인 이미지를 활용할 수 있다. 또는 Fig. 11.과 같이 직접 그린 개념 스케치나 건축 덩어리를 개념 도출 단계에 대응시켜 img2img 기능을 통해 건축 이미지로 변환할 수도 있다. 이와 같은 방법론 자체의 창의적인 변형 및 개선 과정은 현재의 기술적 한계를 넘어 AI 이미지 생성기의 활용 가능성을 확대하기 위한 논의에 긍정적인 영향을 끼칠 수 있을 것이다.

4. 결론

AI 이미지 생성기는 건축 개념 도출 단계에서 보편적인 미에 대한 추상적인 키워드에 대응하여 색감이 뛰어난 창의적인 이미지를 생성하였으며 이들은 건축 초기 단계에서 개념을 표현하기 위한 참고 이미지로 사용될 수 있을 것으로 판단된다. 또한, 개념 이미지를 변형한 건축화 단계의 결과물은 초기 전략의 목적에 부합하여 창의적인 색감 및 형태를 지닌 건축 이미지를 생성하였다.

다만 건축화 단계의 결과물은 일종의 투시도로서 다른 면 혹은 내

부에 대한 정보를 제공하지 않기에 실제 3차원 형태의 구축을 위해서는 상당 부분이 추측에 의존하여야 한다. 그러나 기존의 학습 세트의 영향에서 벗어나기 힘든 일반적인 생성 방식과 달리 사용자의 의도적인 아이디어를 바탕으로 창의적인 건축 이미지로서 기능하는 입체적인 결과물을 생성할 수 있다는 점은 실제 건축 디자인 과정에서 현재의 기술 수준을 바탕으로도 충분히 소통할 수 있는 유용한 도구로 기능할 수 있음을 시사하고 있다.

앞서 언급한 AI 이미지 생성기의 한계는 이미 최근 기초적인 3차원 모델을 형성하는 AI 생성 기술이 개발된 것과 같이 굉장히 빠른 기술 발전 속도와 더불어 이른 시일 내에 해결될 가능성이 대단히 크다. 실제로 본 연구가 시작된 4월 이래로 짧은 수정 기간 사이 등장할 것으로 예측한 기술이 실현된 사례가 상당수 존재한다. 이제는 디자인 과정에서도 컴퓨터가 적극적으로 개입하는 것이 가능해졌으며, CAD와 같이 수요에 의해 프로그램은 더 정교하고 쉽게 사용할 수 있는 방식으로 시장에 등장할 것이다.

AI 이미지 생성기를 통해 단축된 설계 소요 시간과 높은 품질의 결과물을 쉽게 얻어 낼 수 있다는 점은 특히 일정 규모 이상의 기업에서 단순, 디자인 관련 인력에 대한 비용을 줄이도록 유효할 수 있다. 한편 마찬가지로 명성 있는 건축가도 최소한의 인원만으로 본인의 공간에 대한 모델링 혹은 3차원 스캔 자료를 매도 혹은 변형, 발전된 제안으로 생성 및 평가하는 것이 가능해질 것이다. 이는 건축가와 기업의 지적 재산권이 더욱 중요해지며 생산력의 증대와 더불어 전체적인 건축 공간의 품질이 올라가되 핵심 이익이 소수에게 더욱 편중되는 결과를 일으킬 수 있을 것이다.

현 AI의 한계로 인한 낮은 사용성에 망설이기보다 본 연구와 같이 사용을 넘어 시스템을 분석하고 이를 응용할 수 있는 아이디어 및 방법론에 대해 논의하고 제안하는 것은 또 다른 AI 신기술에 대응하고 인간의 건축적인 발상을 새롭게 응용하도록 개발하는 것에 큰 도움이 될 것이다. 이러한 연장선에서 LoRA Model을 통해 직접 AI 이미지 세트를 학습시켜, 이를 통한 결과물과 실제 건축물을 비교하여 효용성을 검증하고 새로운 유형의 건축물로 변환하는 후속 연구를 진행하고자 한다.

Acknowledgement

이 연구는 2023년도 영남대학교 학술연구조성비 지원에 의한 것임(연구과제번호: 223A380053).

References

[1] Midjourney, <https://www.midjourney.com/home/>, 2022.
 [2] R. Rombach et al., High-resolution image synthesis with latent diffusion models, arXiv:2112.10752 [cs.CV], 2022, pp.1-45.
 [3] A. Jaruga-Rozdolska, Artificial intelligence as part of future practices in the architect's work: midjourney generative tool as part of a process of creating an architectural form, *Architectus*, 3(71), 2022.11, pp.95-104.
 [4] J. Ploennings, M. Berger, AI art in architecture, arXiv:2212.09399v1 [cs.AI], 2022, pp.1-21.
 [5] D. Dollens, Stable diffusion, DALL-E 2, Midjourney and metabolic architectures, *Autopoietix*, 2023, pp.1-6.
 [6] M. Radhakrishnan, Is Midjourney-AI the new anti-hero of architectural

imagery & creativity?, *Global Science Journals*, 11(1), 2023.01, pp.94-104.
 [7] 왕택우, 조지영, 주거 및 실내디자인 분야에서 이미지 생성 인공지능 시스템의 현황 및 가능성, 한국주거학회 추계학술발표대회 논문집, 제34권 제2호, 2022, pp.241-242. // (Z.Y. Wang, J.Y. Cho, Status and potentials of image-generator artificial intelligence system in housing and interior design, *Conference Journal of Korean Housing Association*, 34(2), 2022, pp.241-242.)
 [8] 황정석, 조택연, 디자인 인공지능을 활용하기 위한 조형 언어 연구 - 베이비 스키마 조형 언어를 중심으로, 한국공간디자인학회논문집, 제17권 제8호, 2022, pp.327-339. // (J.S. Hwang, T.Y. Cho, A study on the artificial intelligence design controlled by design language - focus on baby schema design language, *Journal of Korean Institute of Spatial Design*, 17(8), 2022, pp.327-339.)
 [9] I. Kant, Critique of the power of judgment, Edition by Paul Guyer, Translation by Paul Guyer and Eric Matthews, New York: Cambridge University Press, 2000, pp.89-127.
 [10] S. Jarvis, "Art, Truth and Ideology" in Adorno: A critical introduction, New York: Routledge, 1998, pp.90-123.

1) 존 폰 노이만(John von Neumann)은 에드박(EDVAC)의 보고서 초안에서 현대 컴퓨터의 기본 구조인 폰 노이만 구조를 설계하였다. 에드윈 제인스(Edwin T. Jaynes)에 의하면 폰 노이만은 기계가 생각하는 것은 불가능하다는 강의를 듣고 자신이 기계가 할 수 없는 것에 대해 정확하게 이야기해 준다면, 자신이 언제든 그것을 수행할 수 있는 기계를 만들 수 있다고 주장하였다. 폰 노이만은 이전에도 모든 처리절차는 컴퓨터에 의해 시뮬레이션 될 수 있다고 주장하였다.
 2) 뮌헨대학교(LMU Munich)의 CompVis 연구실에서 개발한 연구 모델을 기반으로 Stability AI에서 2022년 8월에 배포한 Text-to-Image opensource model이다.
 3) 일반적인 AI 이미지 생성기는 범용성을 위해 총체적인 웹 전반의 이미지 세트를 학습한 경우가 대부분이기에 건축물과 같은 전문적인 결과물을 형성하는 것에 어려움을 겪을 수 있다. 그러나 다음의 이유로 이를 고려하지 않고자 한다. 1) 건축적 이미지를 선별하여 전문적인 모델을 만드는 과정에는 높은 품질의 이미지 학습을 위해서 30GB 이상의 VRam이 요구되는 것으로 알려져 있다. 이는 현시대 기준으로 일반적인 건축가나 학생이 접근하기 힘든 부분이며, 서론에서 언급된바 같이 대중적인 접근성과 유리된 내용은 본 연구의 취지와 맞지 않는다. 그러나 활용성이 높은 소재로써, Low-Rank Adaptation of Large Language Models과 같은 리소스 한계를 극복 가능한 최신 기술을 활용하여 후속 연구에서 다루도록 한다. 2) 비록 건축에 대한 이미지 학습이 전문적인 수준을 만족하지 못할 수도 있으나 이는 절대 건축 이미지를 학습하지 않았다는 의미는 아니며 오히려 건축 이외의 다양한 것들과 더불어 학습 세트에 의한 편향 발생 가능성이 낮은 의미 있는 이미지를 생성할 가능성을 가진다. 특히 클라우드 서비스인 달리 2나 미드저니가 아닌 개인 컴퓨터로 구동 가능한 스테이블 디퓨전을 사용하여 개별적으로 구축된 다양한 학습 모델들로 직접 결과물을 생성하고 적합한 모델을 선정하는 것을 통해 위의 한계를 어느 정도 해소할 수 있을 것으로 예상된다.
 4) Dr. Gloria Willcox의 Feeling wheel은 Robert Plutchik의 Feeling wheel의 영감을 받아 디자인되었으며 본 논문에서는 감정 표현의 범주를 분류하기 위한 목적으로 참고하였다.
 5) 지능이 있어서 질문 답변을 수행할 수 있는 기계가 있어도 그것이 지능을 가졌는지를 튜링 테스트로는 판정할 수 없다는 주장이다.