SCIRES it
SCIentific RESearch and Information Technology
Ricerca Scientifica e Tecnologie dell'Informazione
Vol 15, Issue 2 (2025), 145-162
e-ISSN 2239-4303, DOI 10.2423/i22394303v15n2p145

CASPUR-CIBER Publishing, http://www.sciresit.it

GENERATIVE AI MEETS ARCHITECTURE: TRANSFORMING VISIONS FROM TEXT, SKETCHES, AND 3D MODELS

Matilde Gardini, Cristiana Bartolomei**

*Department of Architecture, Alma Mater Studiorum - University of Bologna, Italy,

Abstract

Generative AI is transforming architectural design, enabling everything from text-based concept generation to sketch animation and hyper-realistic renderings. This study examines whether Gen AI can go beyond visual inspiration to actually generate viable architectural designs. Twenty-one tools were tested using three input types, text prompts, sketches, and 3D models, across specific case studies. A total of 25 renderings were produced and analyzed. The evaluation focused on how well the outputs aligned with the inputs, as well as their structural and functional plausibility. Tools were assessed for their creative flexibility, fidelity to prompts, and architectural coherence. Findings reveal both the potential and current limitations of Gen AI in architectural practice.

Keywords

Architectural design and representation, AI in Architecture, Generative AI, Image generation, AI rendering

1. *Introduction*

Generative Artificial Intelligence (Gen AI) is rapidly emerging as a key driver of digital transformation across multiple industries, including architecture. In particular, the field of architectural representation, long rooted in manual techniques and complex digital workflows, is undergoing a profound revolution. Architectural representation has always served a dual function: not only as a design tool, but also as a means of communicating spatial intentions, aesthetic values, and cultural narratives. From perspective drawings to photorealistic renderings, images have shaped the architectural imagination and its reception. In this long-standing tradition, (Goldschmidt, 2004, Garagnani & Cattoli 2015), Generative AI introduces a new paradigm, one that automates the image-making process while simultaneously expanding the architect's representational vocabulary. During the last decades, architectural representation and design has evolved through distinct phases: from hand-drawing, to the introduction of Computer-Aided Design (CAD) (Gunn, 2002), and later, to Building Information Modeling (BIM) (Visartsakul & Damrianant, 2023). Today, the advent of Gen AI technologies marks the beginning of a new era, one in which visual automation and algorithmic creativity enable architects to explore innovative forms, accelerate design development, and enhance the quality and immediacy of visual outputs (Leach, 2021). The ability of AI to autonomously generate high-quality images, 3D models, and graphic representations is redefining how architects conceive and communicate their projects. Tools based on neural networks, such as Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), and large language models (LLMs), are increasingly integrated into the design pipeline. These systems can interpret textual descriptions, refine sketches, and produce hyper-realistic renderings, creating new opportunities for experimentation and expression. As a result, the architect's creative toolkit is being reshaped, not only enabling faster iterations and broader aesthetic exploration (As, Pal, & Basu, 2018), but also allowing AI to engage with stylistic interpretation and transformation (Bianconi, Filippucci, Migliosi, & Mommi, 2023). Yet, these advancements raise critical questions. While the debate over whether generative AI can be considered "*creative*" continues, a more domain-specific inquiry emerges: *Can Gen AI meaningfully contribute to architectural visualization and design*? In other words, *can it generate outputs that are not only visually appealing but also structurally plausible and functionally viable*?

Tab. 1: Overview of the case studies, input types, and evaluation criteria

Case Study	Input type	Output format	No. of Tools Tested	Evaluation Criteria Applied
Text to image	prompt	JPEG/PNG image	12	1. Photorealism; 2. Architectural Coherence (including coherence with the prompt) 3. Material Realism; 4. Realism of Light, Shadows, and Reflections 5. Details
Sketch to render	Prompt + sketch	JPEG/PNG image	8	1. Photorealism; 2. Architectural Coherence (including coherence with the prompt and the sketch) 3. Material Realism; 4. Realism of Light, Shadows, and Reflections 5. Details
Rendering 3D models	Prompt + 3D model view	JPEG/PNG image	3	1. Photorealism; 2. Architectural Coherence (including coherence with the prompt and the3D model view) 3. Material Realism; 4. Realism of Light, Shadows, and Reflections 5. Details

This study aims to explore that question through a comparative evaluation of 21 Gen AI tools capable of producing realistic architectural renderings. These tools were tested during the summer/autumn of 2024 across a set of curated case studies using three input modes: text prompts alone, text prompts combined with sketches, and text prompts paired with 3D models. Each tool was applied to generate one or two rendering per case, yielding a total of 26 images for analysis. To facilitate comparison across the three test scenarios, Table 1 provides a concise overview of the input modalities, number of tools tested, output formats, and evaluation criteria applied in each case study. This structured summary enhances the readability of the methodology and highlights the consistency of the analytical framework across experiments.

This research addresses two primary questions. First, *how well do Gen AI tools respond to user inputs, both textual and visual, in generating coherent and contextually appropriate designs*? Second, *can these tools support architectural design processes in a meaningful way, producing results that are not only visually compelling but also structurally and functionally plausible?*

By empirically testing a wide range of tools on standardized case studies, this paper seeks to contribute to the growing discourse on the future of Gen AI in architecture. It offers a structured assessment of how these tools perform across various input types and explores their potential roles in enhancing creativity, accelerating workflows, and reimagining the boundaries of architectural representation and design.

From a disciplinary standpoint, generative AI challenges the epistemology of architectural design, questioning the locus of authorship, the role of precedent, and the nature of design iteration. These tensions align with broader discourses on automation, creativity, and human–machine collaboration, making architectural representation an ideal field for testing emerging hybrid workflows.

The structure of this paper is organized as follows: Section 2 reviews a selection of recent studies, providing context and a comparison with our approach. Section 3 outlines the main Gen AI technologies behind the tools used in the experiments, while Section 4 presents the methodology employed. Sections 5, 6, and 7 show the results obtained using the Text-to-Image, Sketch-to-Render, and 3D-to-Render tools, respectively. Finally, Section 8 concludes the paper, followed by the references list.

2. *Generative AI Technologies for Architectural Representation*

Gen AI is rapidly emerging as a key driver of digital transformation across multiple industries, including architecture. In this context, architectural representation is currently experiencing a fast transformation, moving away from its traditional use of manual drafting and specialized digital tools. The rise of Gen AI signals a new era in which visual automation, data-driven design, and algorithmic creativity empower architects to explore novel forms, streamline workflows, and elevate the immediacy and precision of visual outputs. Central to this transformation are powerful deep learning technologies, including CNNs, GANs, Transformers, and LLMs. These models are reshaping how visual information is processed, interpreted, and synthesized, opening new frontiers in how architectural ideas are conceived and communicated (Leach, 2021, Beyan & Rossy, 2023, Horváth & Pouliou, 2024).

CNNs have been foundational in enabling AI to process architectural imagery. Their ability to extract spatial hierarchies from images makes them effective for tasks such as segmentation, elevation extraction, and detail enhancement. Tools like Adobe Firefly's image editing features (Adobe Firefly, 2025), use CNNs for upscaling and refining design visuals, enabling the conversion of rough drawings into clean, detailed representations.

Fig. 1: Architectural image created via Artbreeder (Artbreeder, 2025)

GAN have revolutionized visualization by generating realistic scenes from basic inputs. GAN-based tools, e.g. Artbreeder (Artbreeder, 2025), produce compelling visual content by learning architectural styles and translating abstract inputs (like semantic maps or sketches) into photo-realistic renderings, as the example in Figure1.

Transformers, originally developed for language tasks, have been adapted into vision models (such as Vision Transformers or ViTs) that excel at interpreting spatial relationships and integrating context. Tools like Midjourney (Midjourney, 2025) use transformer-based models to generate highly detailed images of architecture and artifact, often blending abstract ideas with structural logic in visually coherent ways (Palmieri, 2023).

LLMs, particularly when integrated into multimodal systems, serve as interfaces between human intention and visual output. For instance, tools like OpenAI's ChatGPT (ChatGPT, 2025), combined with DALL·E (DALL-E2, 2025), allow users to generate architectural representations by simply describing their ideas in natural language. These models interpret complex prompts and orchestrate image synthesis accordingly, enabling rapid iteration in early design thinking. An example of image created by using ChatGPT (ChatGPT, 2025) is shown in Figure2.

Fig. 2: Architectural image created via ChatGPT 4.0 (ChatGPT, 2025)

These AI tools are increasingly employed to support the early stages of architectural design and representation, where rapid iteration, conceptual exploration, and visual experimentation are critical. Their ability to generate, refine, and reinterpret visual inputs, such as sketches, textual descriptions, or spatial elements, makes them particularly valuable in preliminary design workflows.

3. *Related work*

This Section reviews selected literature, with particular emphasis on contributions that align with the research questions addressed in this paper. The application of generative AI techniques in the early stages of architectural design and representation has been the focus of a growing body of research. This area is well documented in several recent review articles that offer comprehensive overviews of current methodologies, tools, and challenges. Among these, three key works have been particularly influential in framing this study: Beyan and Rossy (2023), Li, Zhang, Du, Zhang, and Xie (2024), and Rane, Choudhary, and Rane. (2024).

Beyan and Rossy (2023) offers a qualitative overview of AI image generators in architecture, emphasizing their influence, challenges, and potential future applications. The authors focus primarily on text-to-image generation tools like DALL·E (DALL-E2, 2025) and Midjourney, (Midjourney, 2025) analyzing their implications for creativity, accessibility, and user proficiency. A technically comprehensive review of generative AI models, GANs, VAEs, diffusion models, and large-scale foundation models, is presented in Li et al. (2024). They map the application of tools across six architectural design phases, from concept to structural detailing. The paper stands out for its systematic classification of tools by model type and design task, offering a granular understanding of how Gen AI integrates into distinct stages of architectural workflows. An approach similar to the one in Li et al. (2024), was employed in a preliminary paper for this study, which mapped over 80 AI applications to various phases of architectural design and representation (Gardini & Bartolomei, 2025). Finally, Rane et al. (2023) take a broader interdisciplinary perspective, reviewing generative AI, not only in architectural design, but also in engineering, urban planning, construction, and sustainability. This review explores tools like ChatGPT (ChatGpt, 2025) within conceptual, collaborative, and regulatory frameworks, highlighting how LLMs influence ideation, communication, and compliance. All three papers include extensive bibliographies relevant to the use of AI-generated images in the early stages of architectural design and representation.

The remainder of this section focuses specifically on four recent works that examine, propose, or apply methods to assess the alignment of generated images in different applications of AI generated images to architecture.

In Zhang, Fort, and Mateu (2023), authors present a system that leverages Stable Diffusion, fine-tuned on images derived from Gaudí's original manuscripts, to generate realistic visual interpretations from minimal textual and visual prompts. Their goal is to explore how generative AI can replicate or reinterpret Gaudí's unique architectural style. To assess the outputs, they adopt four evaluation dimensions: (1) Perceptual Processing, examining the authenticity and aesthetic appeal of the generated designs; (2) Cognitive Processing, evaluating formal cohesion and compositional harmony; (3) Emotional Response, assessing the affective impact of the images; and (4) Creativity, focusing on the novelty and originality of the results. These dimensions are well-suited to the study's objectives, as they reflect the need to measure not only technical fidelity to Gaudí's visual language but also the psychological and artistic resonance of the generated imagery. However, these dimensions do not address the functional or structural accuracy of the generated images. Moreover, the study presented in Park, Ergan, and Feng, C. (2024) investigates whether advanced generative design models can replicate the implicit architectural heuristics used by professionals when designing residential floor plans. The authors compare AI-generated layouts with those produced by architects using statistical tests, focusing on three key evaluation dimensions: (1) proportional space allocation, examining whether shared spaces (e.g., living rooms) receive more area than private ones (e.g., bathrooms); (2) privacy-based spatial organization, assessing the placement of private spaces to reduce visibility and access; and (3) adjacency logic, evaluating the functional relationship between spaces. While generative AI demonstrates some capacity to mimic these spatial conventions, it often fails to reproduce deeper design logic consistently. The assessment methods are strongly grounded in the type of images generated (floor plans) and are not suitable for evaluating other types of generated images.

Zhou & Xiang (2024) present a study evaluating the usefulness and perceived quality of AI-generated residential layouts by comparing them with human-designed plans. The evaluation involved asking a group of professionals to identify the best layout within a set, to distinguish which designs they believed were AI-generated, and to reflect on whether such generated layouts could serve as sources of design inspiration or be used effectively for concept design communication. Results show that while professionals could sometimes distinguish AI outputs, many AI-generated layouts were still considered visually appealing and suitable for early-stage design tasks. The study highlights the potential of generative models in architectural workflows, while also emphasizing the need for clearer quality benchmarks.

Finally, Albaghajati, Bettaieb, and Malek (2023) explored the integration of text-to-image generative models in architectural design, emphasizing both their creative potential and quality-related challenges. Through interviews with sixteen experienced designers, the study found that while AI-generated images enhanced ideation and visualization in early design stages, significant limitations remain in evaluating and achieving high-quality outputs. Participants cited frequent issues with realism, cultural representation, and prompt construction, attributing these to dataset biases and the "black box" nature of generative models. The study also underscored the subjective and iterative nature of assessing image quality, often requiring multiple prompt refinements and human curation. These findings highlight the need for more robust, context-aware evaluation frameworks to ensure that AI-generated visuals meet the functional and aesthetic standards of professional design practice. These four studies offer distinct methodological approaches to evaluating AI-generated architectural images from the perspective of design professionals. While they provide valuable insights, each focuses on specific applications that differ in scope and objective from this study. As a result, the evaluation dimensions adopted here are informed by these works but are not directly derived from their assessment frameworks. Instead, they have been adapted to suit the particular goals and context of this research.

4. *Methodology*

We conducted our trials on 21 tools selected among 80 GenAI application available on line. Our methodology is based on three standardized case studies, each corresponding to one of the selected input methods: text prompts alone, text prompts combined with sketches, and text prompts combined with 3D models. All case studies use the same prompt: *"Photo-realistic image of a modern, ecologic house, made of concrete and glass, located in a wide garden with a water element, during the afternoon, with people living in the place."*.

We requested a photorealistic rendering to obtain results that resemble real-world architectural imagery. The prompt specifies key elements related to style (modern, ecologic), materials (concrete and glass), and the surrounding environment (a wide garden with a water feature, included to capture reflections). It also defines the time of day, and thus the lighting conditions, as afternoon. The inclusion of people ("*with people living in the place*") is intended to enhance the realism and liveliness of the visualization.

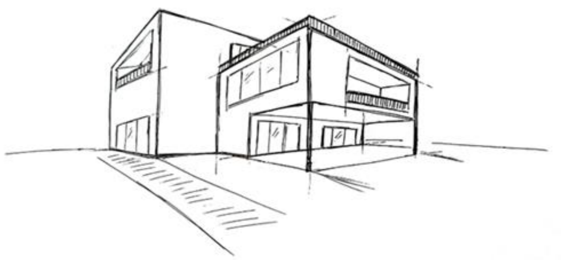

Fig. 3: Sketch used in the sketch-to-render case study

Fig. 4: Sketch used in the sketch-to-render case study (Trimble, 2024).

Figure 3 shows the sketch used in the sketch-to-render case study, while Figure 4 presents the 3D model used in the model-to-render case. The model, downloaded from Trimble's 3D Warehouse (Trimble, 2024), was chosen for its close alignment with the descriptive content of the prompt. Both inputs broadly align with the prompt; however, some key elements, such as the water feature, are missing and are expected to be reflected in the final renderings. Additionally, certain discrepancies are noted: the sketch depicts a structure that appears too thin for a concrete building, and the model includes wood cladding, which contradicts the prompt's specification of concrete and glass. In comparing the tools, traditional technical aspects were not considered (e.g image generation speed, memory or computational requirements). In all cases, image generation was completed within a few seconds.

Instead, the evaluation was informed by key comparative studies in the literature presented in Section 3, which helped identify specific and meaningful criteria relevant to architectural representation and design. These criteria are:

1. *Photorealism*: The system's ability to produce images that closely resemble photographs.
2. *Architectural Coherence*: The system's ability to generate architecture that is (A) *stylistically*, (B) *functionally*, and (C) *structurally* consistent with the given requirements (expressed both trough prompt and image input).
3. *Material Realism*: The system's ability to convincingly reproduce the texture and characteristics of materials, in alignment with what is explicitly requested.
4. *Realism of Light, Shadows, and Reflections*: The overall coherence and believability in the rendering of lighting, shadows, and reflective surfaces within the image.
5. *Details*: The system's accuracy in fulfilling the prompt's specifics and the richness of additional visual elements

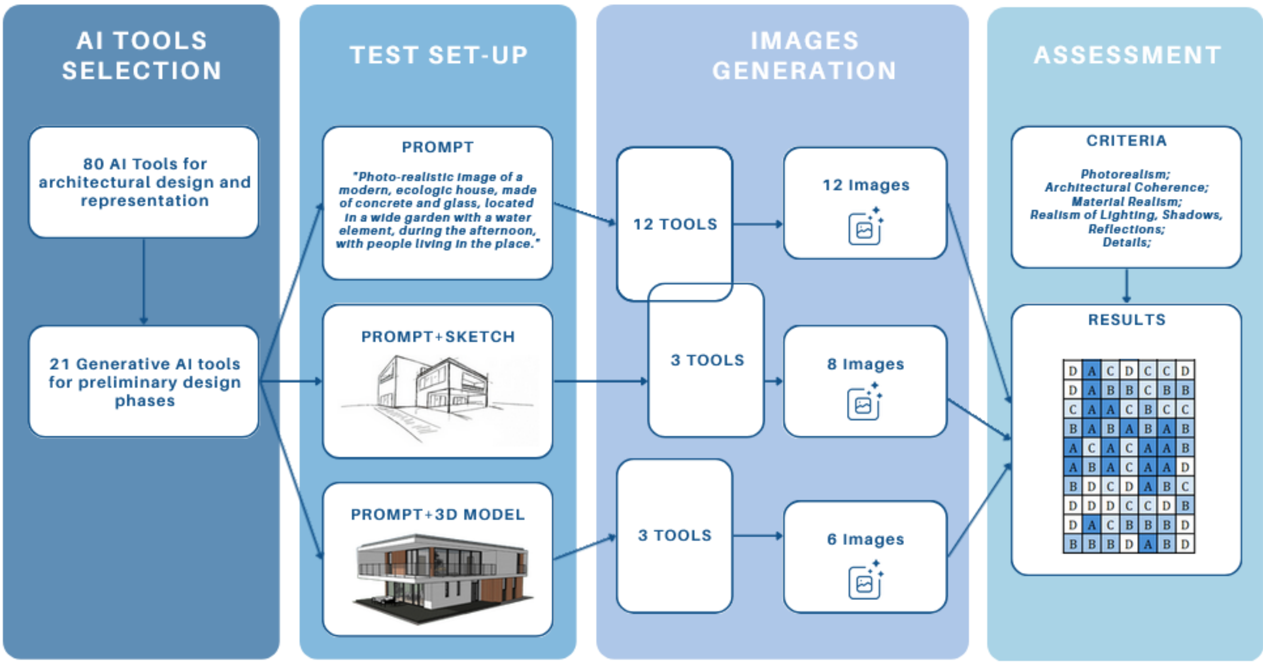

Fig. 4: Four-phase methodological framework comprising selection, testing, generation, and assessment

The four-phase methodological workflow, including tool selection, case study setup, image generation, and assessment based on architectural representation criteria, is shown in Figure 4.In the following sections, we present the set of 25 trials conducted using 21 tools, aimed at addressing the research question and assessing whether these tools meaningfully support architectural design processes, producing results that are not only visually compelling but also structurally and functionally plausible.

5. *Text to Image*

This paragraph presents the evidence and comparisons of the text-to-image generation systems introduced in the previous Section. The wide range of applications and the growing availability of open-source engines have, over time, significantly increased the number of online tools. In this study, conducted during the summer and autumn of 2024, 12 systems were tested and compared, without aiming for exhaustiveness, given the vast array of alternatives available. These tools were selected either because they are part of broader suites for image or digital document processing (e.g., Adobe Firefly (Adobe Firefly, 2025), Bing Image Creator (Bing, 2025), Fotor (Fotor, 2025)); because they function as standalone platforms with large user communities (e.g., Midjourney (Midjourney, 2025), NightCafé (NightCafe, 2025), DreamAI (DreamAI, 2025), Craiyon (Craiyon, 2025)); or because they are specifically geared toward architectural generation (e.g., NovArch AI (Novarch, 2025), LookX AI (LookX AI, 2025), Rendair (Rendair, 2025)). Additionally, some applications based on specific engines were included, such as DALL·E (DALL-E2, 2025), Bing 2025 (Bing, 2025), NightCafè, (NightCafè, 2025) and Stable Diffusion (StableDiffusion 2025) including, DreamStudio (DreamStudio, 2025). The comparison is based solely on the outputs generated from a single prompt input, without the use of optional settings. The results are shown in Figures 5 through 16. The number of images produced varies by tool, from a single image up to 16 alternatives, depending in part on the credit-based payment models.

When multiple images were generated, the one that best matched the prompt from an architectural perspective was selected. Several tools offer a "negative prompt" feature, allowing users to exclude specific elements from the output. Most tools also provide options for adjusting quality and resolution (e.g., aspect ratio or pixel dimensions), the number of images to generate, and stylistic preferences. Two specific cases stand out in relation to the prompt. NovArch (Novarch, 2025) and LookX AI (LookX AI, 2025) include interfaces that help users compose their prompts. As both tools are designed specifically for architecture, users can include terms related to building function, design style, and materials, and may even select reference architects from a curated list of historically significant figures. Craiyon (Craiyon, 2025) goes a step further by not only generating images from the user's prompt but also suggesting additional images based on similar prompts submitted by other users. Craiyon (Craiyon, 2025) goes a step further by not only generating images from the user's prompt but also suggesting additional images based on similar prompts submitted by other users. NightCafé (NightCafe, 2025) is the only tool analyzed that explicitly guarantees SFW (Safe for Work) content, ensuring that the output is suitable for public or professional use by avoiding inappropriate imagery (Merriam-Webster, 2025).

Fig. 6: Image created via MS Bing (Bing, 2025)

Fig. 7: Image created via Adobe Firefly (Adobe Firefly, 2025)

Fig. 8: Image created via DreamAI (DreamAI, 2025)

Regarding training data, Adobe Firefly (Adobe Firefly, 2025) is the only tool that guarantees the use of licensed imagery, either owned by Adobe, covered under Creative Commons licenses, or in the public domain. Some applications also allow the combined use of text prompts and reference images as starting points or sources of inspiration. These features will be examined in the next Section, while the following subsection presents comparison between images generated from prompt.

5.1 Photorealism

Some of the tools tested produce images comparable to high-quality renderings, but still far from being mistaken for real photographs, for example, Bing Image Creator (Bing 2025) (Figure 6). Exceptions include the outputs from Adobe Firefly (Adobe Firefly, 2025) (Figure 7), DreamAI (DreamAI, 2025) (Figure 8), Novarch (Novarch, 2025) (Figure 9), and MidJourney (Midjourney, 2025) (Figure 10), which appear realistic overall. NightCafé AI (NightCafè, 2025) also generates images in which certain elements show sufficiently realistic appearance (Figure 11).

Fig. 9: Image created via Novarch (Novarch, 2025)

Fig. 10: Image created via MidJourney (Midjourney, 2025)

Fig. 11: Image created via Nightcafè (NightCafè, 2025)

Fig. 12: Image created via DreamStudio (DreamStudio, 2025)

Fig. 13: Image created via Craiyon (Craiyon 2025)

5.2 Architectural Coherence

A. *Stylistic Coherence*: All systems correctly interpreted the prompt's request to generate "modern" architecture". In some cases, references to specific architects are evident, such as Mies van der Rohe in Novarch (Novarch, 2025) (Figure 9), and Le Corbusier in both DreamStudio AI (DreamStudio, 2025) (Figure 12) and LookX AI (LookX AI, 2025) (Figure 16).

B. *Functional Coherence*: Functionally, all generated buildings align with the prompt's requirement to depict a house. However, the explicit request for an "ecologic" building was largely overlooked. None of the proposals visibly include insulation (concrete is typically exposed) or solar and/or photovoltaic panels. The only "green" solutions came from MidJourney (Midjourney, 2025), which includes hints of brise-soleil, and LookX AI (LookX AI, 2025), which features a green roof, respectively shown in Figures 11 and 16. The term "*ecologic*" did influence the surrounding context: many gardens are depicted with dense vegetation. In some cases, this extended to interior spaces, with visible plants inside the buildings, e.g Bing Image Creator (Bing, 2025) (Figure 6) and Stable Diffusion AI (Stable Diffusion, 2025) (Figure 14). However, in many cases, water features appear either decorative or disconnected from the building's functional layout, e.g., Craiyon (Craiyon, 2025) (Figure 13) and DreamAI (DreamAI, 2025) (Figure 8).

Fig.14: Image created via Stable Diffusion (StableDiffusion, 2025)

C. Structural Coherence: many images depict buildings with disproportionately large overhangs relative to the supporting structures, especially considering the specified use of concrete. This issue is particularly noticeable in outputs by Adobe Firefly (Adobe Firefly, 2025) (Figure 7), NightCafé AI (NightCafè, 2025) (Figure 11), and Stable Diffusion (StableDiffusion, 2025) (Figure 14). There is generally a lack of internal and external pillars, or the ones present are too few or too slender to plausibly support upper structures. These shortcomings may stem from an effort to replicate a "modern" aesthetic but result in structurally implausible designs. Only the outputs from Novarch AI (Novarch, 2025), Craiyon (Craiyon, 2025), and Rendair (Rendarir, 2025) (respectively Figures 9, 13, 17), which feature single-story buildings, demonstrate structural consistency. In other cases, while the structures might be technically feasible, their articulation remains unclear

Fig.15: Image created via Fotor (Fotor, 2025)

Fig. 16: Image created via LookX AI (LookX AI, 2025)

5.3 Material Realism

With the exception of Craiyon (Craiyon, 2025) (Figure 13), where the use of concrete is not clearly evident, all other tools generated images prominently featuring the two requested materials: concrete and glass. In several cases, particularly Bing Image Creator (Bing, 2025) (Figure 6) and NightCafé AI (NightCafè, 2025) (Figure 11), the textures and rendering of materials appear especially realistic.

5. 4 Realism of Lighting, Shadows, and Reflections

The prompt explicitly specified an afternoon setting, which all tools respected except Craiyon (Craiyon, 2025) (Figure 13), which shows a dusk scene. Reflections on glass are generally convincing, though some tools struggle with transitions between reflective surfaces and interior spaces, as seen in Stable Diffusion (StableDiffusion, 2025) (Figure 14). Reflections in water elements are typically less precise and often exhibit distorted perspective; Craiyon (Craiyon, 2025) (Figure 13) offers the most evident example of this flaw.

Fig. 17: Image created via Rendair (Rendair, 2025)

Fig. 18: Imprecise details created by (a) Adobe Firefly (Adobe Firefly, 2025) and (b) DreamStudio (DreamStudio, 2025).

5.5 Details

A specific detail in the prompt, the presence of people, was included in only 6 of the 12 tools: Bing Image Creator (Bing, 2025), Adobe Firefly (Adobe Firefly, 2025), MidJourney (MidJourney, 2025), Stable Diffusion (StableDiffusion 2025), Fotor (Fotor, 2025), and LookX AI (LookX AI, 2025) (respectively in Figures 6, 7, 11, 14, 15, 16). These human figures, although present, are often poorly rendered. When zoomed in, they reveal deformations and unrealistic proportions (Figure 18.a). All generated images include some level of furnishing, both indoor and outdoor. However, these elements are often inaccurate or stylized in ways that, when enlarged, display unrealistic forms or details (Figure 18.b). These limitations in detail are common in AI-generated imagery (Campo & Leach, 2022) and become especially apparent at larger viewing sizes.

5.6 Final Remarks

To summarize, the tools that delivered the most suitable results for architectural representation and design purposes were Novarch (Novarch, 2025) (Figure 9) and Rendair (Rendair, 2025) (Figure 17), both of which were specifically developed for architectural use. The rigidity of the test (applying the same prompt across all tools without prompt engineering or parameter adjustments) represents a clear limitation of the study. However, this uniform approach enabled a direct comparison of each system's baseline performance through straightforward, unmodified usage. The process of testing, cataloguing, and evaluating the results helped illuminate the core strengths and limitations of each application.

6. *Sketch to render*

As for prompt-to-image generation systems, there is a wide range of online applications that generate images from sketches or inspirational images. In this study, eight of such tools were tested and compared, without claiming to be exhaustive. Among these, two (LookX AI (LookX AI, 2025) and Rendair (Randair, 2025)) were examined in greater detail in the tests presented in Section 5. In addition, six other applications were analyzed, selected based on their availability as free or trial versions at the time of testing and their relevance to architectural design and representation. These include Architect Render (Architect Render, 2025), Gaia (Gaia, 2025), Mnml.ai (Mnml, 2025), PromeAI (PromeAI, 2025), ReRenderAI (ReRender, 2025), and Visoid (Visoid, 2025). The results are shown in Figures 18 through 25.

Analogously to Section 5, the following subsection presents comparison between images generated from sketch and prompt.

Fig. 19: Image generated via Rendair (Rendair, 2025)

Fig. 20: Images generated via Architect Render (ArchitectRender, 2025)

6.1 Photorealism

Half of the tools analyzed demonstrated a good level of photorealism, i.e. generate images that resemble real photographs. In particular, the images produced by Rendair (Rendair, 2025) and Architect Render (ArchitectRender, 2025), respectively shown in Figure 19 and 20 include highly realistic elements. In contrast, the outputs from Gaia (Gaia, 2025), Visoid (Visoid, 2025), and PromeAI (PromeAi, 2025), appear more illustrative and resemble sketches rather than realistic renderings (respectively in Figures 22, 23, and 25).

6.2 Architectural Coherence

A. Coherence with the Sketch: Unlike the style evaluation in Section 5, in this test the sketch defines the intended style. Thus, the analysis focused on fidelity to the sketch rather than on stylistic aspects. All tools produced images that closely matched the input sketch. The most noticeable discrepancies involved are:

Fig. 21: Image generated via ReRenderAI (ReRender, 2025)

- *Additional volumes*: Some tools introduced architectural elements not present in the original sketch. For example, in ReRenderAI (ReRender, 2025) (Figure 21) and Gaia (Gaia, 2025) (Figure 22), balconies were transformed into terraces, and Gaia's output even included an added volume behind the building. Several images also show extra concrete features like perimeter walls or canopies.
- *Altered railings, recesses, and windows*: These details were often ignored or only partially rendered. In particular, Gaia (Gaia, 2025) (Figure 22) failed to depict recessed areas, replacing them with terraces and windows. A second-floor door was transformed into a window, and railings were omitted entirely. This omission was also noted in outputs from ReRenderAI (ReRender, 2025) (Figure 21) and Rendair (Rendair, 2025) (Figure 19).

B. *Functional Coherence*: All images represent single-family homes, in line with the input sketch. However, the explicit request for an "*eco-friendly*" building used in the prompt was ignored by all tools. None of the renderings depicted insulation, solar panels, or passive shading systems. These elements were not in the sketch, and including them in the prompt was not sufficient for their generation. All outputs included large gardens and at least one water element as requested, although the water features were often ambiguous: lakes and pools were sometimes indistinguishable, and access to them was unclear, as seen in Visoid (Visoid, 2025) (Figure 23).

Another design inconsistency was the central access road, which was variously interpreted as a pool in Rendair (Rendair, 2025) (Figure 19) or as a pier in Architect Render (ArchitectRender, 2025) (Figure 20). In Mnml (Mnml, 2025) (Figure 24), this element lost its intended function altogether. In general, external elements, described only in the prompt and not depicted in the sketch, were poorly integrated, often appearing disjointed, poorly arranged, or inaccessible.

Fig. 22: Image generated via Gaia (Gaia, 2025)

Fig. 23: Image generated via Visoid (Visoid, 2025)

Fig. 24: Image generated via Mnml.ai (Mnml, 2025)

Fig. 25: Image generated via PromeAI (PromeAI, 2025)

Fig. 26: Image generated via LookX AI (LookX AI, 2025)

C. *Structural Coherence*: Several tools generated structures with questionable static stability, mainly due to unrealistic representations of the pillars supporting large cantilevers. While the sketch implied a steel structure with slender supports, the prompt specifically requested the use of concrete and glass. Many tools misinterpreted this, resulting in pillars that were too thin to bear the load (e.g., LookX AI (LookX AI, 2025), (Figure 26), or eliminated the supports entirely in favor of full glazing, as seen in Mnml.ai (Mnml, 2025) (Figure 24) and ReRenderAI (ReRender, 2025) (Figure 21). In two cases. Rendair (Rendair, 2025) (Figure 19) and Gaia (Gaia, 2025) (Figure 22), concrete pillars or support walls were added, improving stability but introducing asymmetries. The tools that best interpreted the mix of materials and produced structurally sound architecture were Visoid (Visoid, 2025) (Figure 23) and PromeAI (PromeAI, 2025) (Figure 25).

Fig. 27: Imprecise details generated by (a) Gaia (Gaia, 2025) and (b) LookX AI (LookX AI, 2025).

6.3 Material Realism

All tools demonstrated a high level of material realism, i.e. they represent textures and physical properties credibly. Materials such as concrete and glass were rendered convincingly, and steel was depicted appropriately where it could be inferred from the sketch.

6.4 Realism of Lighting, Shadows, and Reflections

All tools correctly interpreted the lighting conditions as late afternoon. Light behavior in the scenes was coherent, and reflections on glass and water were generally realistic. Exceptions include LookX AI (LookX AI 2025) (Figure 26) and Gaia (Gaia, 2025) (Figure 22), where glass reflections appeared less convincing.

6.5 Details

Most tools included few of the specific details requested in the prompt beyond what was shown in the sketch. Only LookX AI (LookX AI, 2025) (Figure 26) and Gaia (Gaia, 2025) (Figure 22) depicted human figures. Furnishings and vegetation were also less abundant than in the images discussed in Section 5, which were generated using prompts alone. As a result, fewer visual inconsistencies typically seen in AI-generated images were observed (Campo & Leach, 2022). Two of them are depicted in Figures 27.a and 27.b. Finally, some tools added borders to adapt the render format to the sketch, such as Rendair (Rendair, 2025) (Figure 19) and Visoid (Visoid, 2025) (Figure 23).

6.6 Final Remarks

The tools that produced the most suitable outputs for architectural representation and design purposes were Rendair (Rendair, 2025) (Figure 19) and Architect Render (ArchitectRender, 2025) (Figure 20). In both cases, the renderings accurately reflected both the sketch and the prompt, generating structurally sound designs while including many of the relevant architectural details.

7. *Rendering 3D models*

This section is devoted to AI tools used to generate renderings from two- or three-dimensional models, integrated directly with professional design software. These systems work on a single model view at a time, applying image-generation technologies to a highly detailed base image derived from the 3D model. To ensure consistent and comparable testing conditions, the evaluation focused on AI based rendering tools that integrate directly with the SketchUp modeling environment (Sketchup, 2025), a widely used platform in architectural design. Among the five tools identified with this compatibility, three were selected for in-depth analysis: ArkoAI (ArkoAI, 2025) SketchUp Diffusion (SketchUp Diffusion, 2025), and Veras (Veras, 2025). These tools were chosen for their ability to incorporate AI-driven rendering features within the SketchUp workflow. Two additional tools, D5Render (D5Render, 2025) and Chaos Enscape (Chaos Enscape, 2025), were excluded from testing due to specific limitations: the former required system resources beyond the available configuration, while the latter did not include AI features in its trial version at the time of the test.

Fig. 28: Image generated via ArkoAI (ArkoAI, 2025).

Fig. 29: Image generated via ArkoAI (ArkoAI, 2025).

The following subsection presents comparison between results of these three tools

7.1 Photorealism

The tool ArkoAI (ArkoAI, 2025), which focuses exclusively on rendering 3D structures without incorporating their environmental context, produced less photorealistic results (Figures 28 and 29). However, this simplified approach could be functional when overlaying the rendering onto real photographs of the project site. When evaluating the 3D structure alone, the result is photorealistic. In contrast, SketchUp Diffusion (SketchUp Diffusion, 2025) and Veras (Veras, 2025) yielded less convincing levels of realism, with variability depending on the specific image and its components (Figures 30, 31, 32 and 33 respectively).

Fig. 30: Image generated via SketchUp Diffusion (SketchUp Diffusion, 2025).

Fig. 31: Image generated via SketchUp Diffusion (SketchUp Diffusion, 2025).

7.2 Architectural Coherence

A. *Coherence with the 3D Model*: All three tools preserved the structural form of the model without altering its volumes. However, differences were noted in the placement and design of openings (e.g., windows and doors). One example is Sketchp Diffusion (SketchUp Diffusion, 2025) (Figures 30 and 31), where window configurations deviate from the original. In tools that allowed geometric fidelity adjustment, SketchUp Diffusion (SketchUp Diffusion, 2025) and Veras (Veras, 2025), maximum alignment was selected.

Fig. 32: Image generated via Veras (Veras, 2025).

Fig. 33: Image generated via Veras (Veras, 2025).

B. *Functional Coherence:* All renderings correctly represent single-family houses, consistent with the source model. However, functional elements not included in the model, such as pathways and surrounding green spaces, were either omitted or implemented poorly. In Veras (Veras, 2025) (Figure 33), for instance, the house is entirely surrounded by water. Other external features, requested in the prompt but not modeled, introduced additional issues, such as unclear access points to pools or disjointed material transitions in the garden. These are illustrated in Figure 30 and 31, generated by SketchUp Diffusion (SketchUp Diffusion, 2025). The prompt's reference to an "ecologic" house was generally ignored, with the exception of Veras (Veras, 2025), which added green roofs (32 and 33). In ArkoAI (ArkoAI, 2025) (Figure 28 and 29), the absence of background or environmental modeling negatively affected the rendering of external elements requested in the prompt, e.g. the lawn awkwardly placed beneath the cantilever in Figure 29.

C. *Structural Coherence*: Fidelity to the original model ensured that all rendered buildings were structurally plausible. Minor structural issues were noted in one rendering by Veras (Veras, 2025) (Figure 33), where some rear load-bearing walls appeared transparent.

Fig. 34: Inaccurate details produced by ArkoAI (ArkoAI, 2025), and Veras (Veras, 2025).

7.3 Material Realism

Both SketchUp Diffusion (SketchUp Diffusion, 2025) and Veras (Veras, 2025) correctly interpreted the materials specified in the prompt and applied them appropriately in the rendering. The level of detail included realistic textures, even down to formwork patterns on concrete elements (Figure 30, 31, 32, and 33). ArkoAI (ArkoAI, 2025) presents plastered surfaces that may have resulted from ignoring the prompt instructions to use concrete and glass, instead faithfully reproducing the model without incorporating the textual input. We considered it unlikely that the plastering was intended to cover insulation, so this option was not evaluated as ecological (Figures 28 and 29).

7.4 Realism of Lighting, Shadows, and Reflections

All rendering outputs featured relatively accurate shadows, as they were derived directly from the model. However, reflections were notably less realistic, especially on glass surfaces. Water reflections fared somewhat better in Veras (Veras, 2025) (Figures 33), some reflections blended with interior details or displayed content inconsistent with the surrounding exterior. Afternoon lighting, as requested in the prompt, was correctly applied in all cases. Except for Figure 33 by Veras (Veras, 2025), none of the renderings included artificial interior lighting, making indoor areas appear overly dark and indistinct; for example, in Figure 30 by, SketchUp Diffusion (SketchUp Diffusion, 2025).

7.5 Details

All systems ignored the prompt's request to include people in the renderings. Furthermore, the few outdoor furnishings and other fine details are often marred by typical AI-generation flaws (Campo & Leach, 2022), some of which are shown in Figure 34.

7.6 Final Remarks

Only three tools were evaluated in the 3D model-to-render category, and overall results were moderately unsatisfactory. The most effective tool in this test was SketchUp Diffusion (SketchUp Diffusion, 2025)**,** which also benefits from being integrated into SketchUp license. The 3D model (Trimble, 2025) used details materials that differed from those specified in the prompt, a discrepancy that the tools did not always manage appropriately. Future developments in this line of research may include: (1) tests using models with varying levels of detail, including those with no predefined materials; (2) further tests based on screenshots of the SketchUp model processed using the tools evaluated in Section 5.

8. *Conclusions and future work*

This paper presents a four-phase methodology for evaluating AI-generated images from the perspective of architectural representation and design. The approach, illustrated in Figure 4, was tested through a first set of experiments assessing the capabilities of current generative tools.

The experimental results highlight a clear relationship between the explicitness of the input and the architectural coherence of the output. As summarized in Table 2, text to image cases (Figures 6–17) concentrate C–D grades, especially on structural and functional coherence, and perform unevenly on material and lighting realism and detail quality. Sketch-conditioned cases (Figures 19–26) shift toward predominantly B outcomes, recovering spatial and structural logic while stabilizing material and light realism. Despite their limited spatial and structural consistency for design development, text to image results proved generally faithful to the requested style and valuable as sources of visual inspiration. Model-conditioned cases (Figures 28–33) achieve consistent B–A ratings across architectural coherence and show the most reliable material and lighting realism, indicating that visual grounding restores design control and improves overall architectural consistency.

Table 2: Assessment of the 26 analyzed images based on five criteria and corresponding subcriteria, using a four-level rating scale (A = excellent, B = good, C = fair, D = poor).

	1	2.A	2.B	2.C	3	4	5
Figure 6	D	B	B	B	A	B	B
Figure 7	A	A	A	C	B	A	C
Figure 8	A	C	D	C	B	C	B
Figure 9	A	A	B	A	B	A	B
Figure 10	A	A	A	D	C	C	C
Figure 11	B	B	D	D	A	C	B
Figure 12	D	A	C	D	C	B	C
Figure 13	D	A	D	A	D	D	D
Figure 14	D	A	C	D	C	C	D
Figure 15	D	A	B	B	C	B	B
Figure 16	C	A	A	C	B	C	C
Figure 17	B	A	B	A	B	A	B
Figure 19	A	C	A	C	A	A	B
Figure 20	A	B	A	C	A	A	D
Figure 21	B	D	C	D	A	B	C
Figure 22	D	D	D	C	C	D	B
Figure 23	D	A	C	B	B	B	D
Figure 24	B	B	B	D	A	B	D
Figure 25	C	A	B	B	C	C	B
Figure 26	D	D	D	D	B	D	C
Figure 28	C	A	B	A	B	B	C
Figure 29	C	A	B	A	B	B	D
Figure 30	D	B	C	A	A	C	C
Figure 31	B	B	C	A	A	B	C
Figure 32	C	A	A	B	A	D	C
Figure 33	D	B	B	B	A	C	C

Beyond specific performance patterns, the study highlights the central role of prompt design and user interaction in shaping the architectural quality of AI-generated outputs. The act of defining prompts, selecting visual references, and curating constraints configures a new operative field in which the architect's authorship evolves from drawing to the orchestration of generative processes. Within this framework, interaction itself becomes a form of design, a space where choices and iterations acquire the same value once attributed to the manual act of drawing.

This dependency on input quality and interpretation also exposes the lack of shared benchmarks for evaluating AI-driven images, calling for reproducible and transparent criteria within the disciplinary domain of architectural representation.

In this perspective, drawing regains a renewed strategic and cultural role—not only as a representational product but as a reflective instrument mediating between human intention, visual imagination, and the machine's generative logic. The volatility of generative tools limits the permanence of specific results, yet this instability confirms the need for methods able to adapt to technological change. The research value lies not in the durability of findings but in framing the evolution of AI-driven representation within the dynamics of architectural design.

As generative systems enter professional and educational workflows, architects will increasingly face questions of authorship, agency, and accountability. Human creativity now operates within datasets and shared visual memories, engaging collective cultural materials and conventions. This condition challenges traditional notions of originality and expands the ethical horizon of drawing: designing with AI means negotiating between personal intent, visual heritage, and algorithmic logic. The capacity to guide and interpret machine-generated representations thus becomes a key aspect of professional ethics, extending design responsibility into new technological and cultural territories.

Future research should refine the proposed methodology toward more systematic and reproducible evaluations, involving expert panels and quantitative scoring. Such developments could support shared quality benchmarks for generative AI in architecture and strengthen the methodological foundations of this emerging field.

In the end, the integration of generative AI into architectural representation does not replace the act of drawing but extends it, turning every image into a site where technology and imagination meet to reimagine the future of design. Research in this field should continue to guide and critically shape this integration, ensuring it evolves from experimentation into a conscious and structured component of the design process.

REFERENCES

Albaghajati, Z. M., Bettaieb, D. M., & Malek, R. B. (2023). Exploring text-to-image application in architectural design: Insights and implications. *Architecture, Structures and Construction*, *3*(4), 475-497.

Adobe. (2025). *Adobe Firefly*. https://firefly.adobe.com/

Artbreeder. (2025). *AI image generator*. https://pollo.ai/ai-image-generator

Architect Render. (2025). *[Homepage]*. https://www.architectrender.com/

ArkoAI. (2025). *ArkoAI for Sketchup*. https://arko.ai/sketchup

As, I., Pal, S., & Basu, P. (2018). Artificial intelligence in architecture: Generating conceptual design via deep learning. *International Journal of Architectural Computing, 16*(4), 306-327.

Beyan, E. V. P., & Rossy, A. G. C. (2023). A review of AI image generator: influences, challenges, and future prospects for architectural field. *Journal of Artificial Intelligence in Architecture*, *2*(1), 53-65.

Bianconi, F., Filippucci, M., Migliosi, A., & Mommi, C. (2024). AI-D. Algorithmic Experiments with AI to Aid Drawing. *SCIRES-IT – SCIentific RESearch and Information Technology*, *14*(2), 31–48. http://dx.doi.org/10.2423/i22394303v14n2p31

Bing Creator. (2025). https://www.bing.com/images/create

Chaos Enscape. (2025). https://www.chaos.com/it/enscape

Campo, M., & Leach, N. (Eds.). (2022). *Machine hallucinations: architecture and artificial intelligence*. John Wiley & Sons.

ChatGPT. (2025). *ChatGPT*. https://chatgpt.com/

Craiyon. (2025).*Craiyon*. https://www.craiyon.com/

D5render. (2025). *D5 Render*. https://www.d5render.com/

DreamAI. (2025). *DreamAI*. https://dream.ai/

Fotor. (2025). *AI image generator*. https://www.fotor.com/ai-image-generator/

Gaia. (2025). *Gaia*. https://gaia.gs/

Garagnani, S., & Cattoli, G. (2015). Virtual modeling for architecture: Teaching today digital representation techniques to tomorrow's professional designers. *SCIRES-IT - Scientific Research and Information Technology*, *5*(2), 15–26. http://dx.doi.org/10.2423/i22394303v5n1p59

Gardini, M., & Bartolomei, C. (2025). *From concept to reality: Analyzing the role of AI in architectural visualization and design.* In Proceedings of the 28th EVA Conference Berlin. Fraunhofer Institute.

Goldschmidt, G. (2004). *Design representation: Private process, public image*. In G. Goldschmidt & W. L. Porter (Eds.), Design representation (pp. 203–217). London, England: Springer.

Gunn, W. (2002). *The Social and Environmental Impact of Incorporating Computer Aided Design Technologies into an Architectural Design Process* [Phd]. University of Manchester.

Horváth, A. S., & Pouliou, P. (2024). AI for conceptual architecture: Reflections on designing with text-to-text, text-to-image, and image-to-image generators. *Frontiers of Architectural Research, 13*(3), 593-612.

Leach, N. (2021). *Architecture in the age of artificial intelligence: An introduction for architects*, Bloomsbury Visual Arts.

Li, C., Zhang, T., Du, X., Zhang, Y., & Xie, H. (2024). Generative AI for architectural design: A literature review. arXiv e-prints, arXiv-240

LookX AI. (2025). *LookX AI*. https://www.lookx.ai/

Merriam-Webster. (2025). NSFW, not safe for work. In *Merriam-Webster.com dictionary.* https://www.merriam-webster.com/dictionary/NSFW

Midjourney. (2025). *Midjourney*. https://www.midjourney.com/home

Mnml. (2025). *Mnml*. https://mnml.ai/

NightCafè. (2025). *AI art generator.* NightCafè Creator. https://creator.nightcafe.studio/

Novarch AI. (2025). *Novarch AI*. https://ai.novarch.us/

OpenAI. (2022). *DALL·E 2*. https://openai.com/index/dall-e-2/

Palmieri, A. (2023). Midjourney experimentation: representing nature on a macro scale. *SCIRES-IT – SCIentific RESearch and Information Technology*, *13*(1), 181–188. http://dx.doi.org/10.2423/i22394303v13n1p181

Park, K., Ergan, S., & Feng, C. (2024). Quality assessment of residential layout designs generated by relational Generative Adversarial Networks (GANs). *Automation in Construction*, 158, 105243.

PromeAI. (2025.). *AI design workflow generator*. PromeAI. https://www.promeai.pro/

Rane, N., Choudhary, S., & Rane, J. (2023). Leading-edge technologies for architectural design: a comprehensive review. Available at SSRN 4637891.

Rendair. (2025). *Crea, Modifica, Migliora e Anima—istantaneamente con l'IA*. Rendair AI. https://www.rendair.ai/it/

ReRender. (2025). *[Homepage]*. https://rerenderai.com/

Sketchup. (2025). *Sketchup*. https://www.sketchup.com/en

SketchUp. (2025). *SketchUp Diffusion*. https://help.sketchup.com/it/sketchup-labs/sketchup-diffusion

Stability AI. (2025).*DreamStudio*. Stability AI. https://dreamstudio.stability.ai

Stable Diffusion at Openart. (2025). *Stable Diffusion at Openart*. https://openart.ai/

Visartsakul, B., & Damrianant, J. (2023). A review of building information modeling and simulation as virtual representations under the digital twin concept. *Engineering Journal*, 27(1), 11-27.

Trimble 3DwareHouse. (2024). https://3dwarehouse.sketchup.com/model/5fd976ac-d07b-4fac-8dd2-8947d6fa51a1/Modern-House

Veras (2025). https://www.evolvelab.io/veras

Visoid. (2025) https://www.visoid.com/

Zhang, Z., Fort, J. M., & Mateu, L. G. (2023). Exploring the potential of artificial intelligence as a tool for architectural design: A perception study using Gaudi's works. *Buildings*, *13*(7), 1863.

Zhou, H., & Xiang, S. (2024). Applicability Evaluation and Reflection on Artificial Intelligence-based "Image to Image" Generation of Landscape Architecture Masterplans. *Landscape Architecture Frontiers*, *12*(2), 58-67.