Argumentation Based Systems Journal

www.absj.org
eISSN: 3120-1024

A Generative AI–Based Decision-Support Framework for Early-Stage Architectural Design: Evidence from Sustainable Tiny Houses

Tuğçe Çelik[1,]

[1] Faculty of Architecture and Design, Ostim Technical University, Ankara, Turkey

ARTICLE INFO

Article history:
Received 29 January 2026
Received in revised form 1 March 2026
Accepted 9 March 2026
Available online 19 March 2026

Keywords:

Generative Artificial Intelligence, Design Decision Support, Text-to-Image Models, Architectural Design Reasoning, Sustainable Tiny House

ABSTRACT

This study investigates the potential of generative artificial intelligence–based text-to-image models as early-stage decision-support tools in sustainable tiny house design. Within the conceptual design process, the study systematically models key components—formal schema development, material selection, and spatial atmosphere—through a controlled prompt framework that enables structured comparison of design alternatives. Design outputs generated by diffusion-based models, including ChatGPT (DALL·E 3), Copilot, and Gemini AI, are evaluated using a multi-criteria framework encompassing sustainability principles (use of natural materials, compact planning, energy-efficient openings), typological consistency, and overall design quality. To examine how prompt variations shape design outcomes, three analytical dimensions are defined: (1) morphological decisions (plan layout, roof typology, form), (2) material decisions (timber, recycled composites, hybrid systems), and (3) spatial atmosphere (daylighting, color palette, interior warmth). The findings show that text-to-image workflows not only accelerate the generation of design alternatives but also enable a structured interpretation of how different design options align with specified criteria. In this sense, generated outputs function as interpretable design propositions that can be comparatively assessed within a transparent evaluation logic. The results further suggest that generative AI supports designers by externalizing design reasoning, facilitating the exploration, comparison, and justification of early-stage decisions. These results position generative AI as a viable component of next-generation design decision-support systems, particularly in contexts where rapid iteration, multi-criteria evaluation, and the articulation of design reasoning are critical.

1. Introduction

The early stages of architectural design encompass the most critical decision-making processes, during which designers construct formal schemas, evaluate material options, and define the spatial atmosphere. These phases are characterized by high levels of uncertainty, an abundance of possible alternatives, and considerable pressure on designers to generate well-reasoned decisions. In recent years, generative artificial intelligence (Generative AI) models, particularly diffusion-based text-to-image systems, have begun to transform architectural design practices by offering new possibilities for supporting these early decision phases. Models such as ChatGPT (DALL·E 3), Copilot, and Gemini

[1] *tugce.celik@ostimteknik.edu.tr*
doi: https://doi.org/10.59543/nnc2a471

AI rapidly generate numerous design alternatives, accelerating visualization workflows and providing designers with an expanded space for conceptual exploration.

These technological developments necessitate a reconsideration of intelligent decision-support systems within the field of architecture. The text-to-image workflow functions not merely as a visual output tool but as a computational support mechanism that offers rapid feedback, enables comparative assessment of material and form variations, and allows sustainability criteria to be examined at the earliest stages of design. In the context of tiny house design, this is particularly significant, as parameters such as sustainability, compact planning, energy efficiency, and the use of natural materials must be addressed from the very first steps of the design process.

This study investigates the potential of text-to-image models as decision-support tools in sustainable tiny house design and examines how controlled prompt sets influence the decision-making process across formal, material, and atmospheric variations. In doing so, the research provides a holistic framework for understanding how generative artificial intelligence can evolve into a systematic, evaluable, and reproducible knowledge-production mechanism within architectural design workflows.

2. Literature Review

2.1 The Evolution of Generative Artificial Intelligence and Diffusion Models

Generative artificial intelligence has undergone a significant transformation over the past decade, driven by advances in deep learning architectures. The era of high-quality image generation, initially catalyzed by Generative Adversarial Networks (GANs), evolved into a new phase following the rise of diffusion models after 2020. The Denoising Diffusion Probabilistic Models (DDPM) introduced by Ho et al. (2020) demonstrated that iterative denoising procedures could reliably model complex image distributions, offering higher resolution, more controlled outputs, and fewer artifacts compared to GAN-based approaches (Ho et.al., 2020).

The success of diffusion models paved the way for advanced text-to-image systems such as Stable Diffusion (Rombach et al., 2022), DALL·E 2 (Ramesh et al., 2022), Midjourney, and Imagen. Leveraging latent space representations and CLIP-based text–image alignment mechanisms, these models enable users to generate photorealistic, stylized, or highly conceptual visuals directly from textual descriptions. In particular, latent diffusion models have reduced computational costs while expanding the accessibility and application domains of high-resolution image synthesis.

Furthermore, controllable generation techniques, such as ControlNet, LoRA, and T2I-Adapters, along with methods for style transfer, variation-based design exploration, multi-prompt composition, and hierarchical control of detail levels, have made generative models increasingly suitable for design-oriented workflows. Within fields such as architecture, industrial design, and digital content creation, these models are increasingly regarded as “creative collaborators,” supporting processes of variation generation, formal exploration, and rapid visualization of early design ideas.

This evolution has enabled generative AI models to move beyond mere aesthetic image synthesis tools and become integrated components of decision-making processes, capable of producing multi-scenario outputs and offering context-aware design alternatives, thus positioning them as intelligent design-support mechanisms in contemporary creative practice.

2.2 Text-to-Image Generation and Its Use in Design Disciplines

Text-to-image (T2I) models have become one of the primary driving forces behind the widespread adoption of generative artificial intelligence in creative industries. Initially embraced in fields requiring rapid concept development, such as digital art, fashion, advertising, and game design, these models have quickly evolved into powerful tools for the early stages of design processes. Their ability to transform a textual design intention into multiple visual variations within seconds offers significant advantages for conceptual exploration, form experimentation, style comparison, and visual storytelling.

In recent years, the disciplines of architecture and interior design have also begun integrating T2I models into workflows involving conceptual design, atmosphere creation, material research, and spatial character analysis. Chen et al. (2023) demonstrate that T2I models diversify architectural form exploration and provide designers with an expanded capacity for generating alternatives (Chen et.al., 2023). Gallega & Sumi (2024) similarly highlight that generative models accelerate conceptual discovery in interior environments, particularly in relation to atmosphere, lighting conditions, and material–texture matching (Gallega & Sumi, 2024).

A further notable contribution in this area is presented by Çelik (2025), who developed a three-stage methodology to evaluate diffusion-based T2I-generated housing plans in terms of daylight performance across five climate zones. This method combines T2I generation, redrawing in AutoCAD, and climate-based simulations using Velux Daylight Visualizer. The study reveals that although ChatGPT and Copilot can produce architecturally legible plans, they lack coherent environmental logic with respect to solar orientation, seasonal light variation, and passive daylighting strategies. The findings suggest that while generative models perform well in form production, they have not yet internalized performance-based design knowledge, underscoring the need to strengthen T2I tools as early-stage decision-support components (Çelik, 2025).

Despite their contributions to creative design workflows, T2I models continue to face limitations related to controllability, scale and proportion consistency, spatial logic errors, physical constructability, and material realism. As such, the use of generative models in design disciplines presents both opportunities and methodological responsibilities.

2.3 Early-Stage Decision Making in Architectural Design

Early-stage design decisions are characterized in the literature as processes marked by uncertainty, multiple constraints, limited information, intuitive reasoning, and the need for high creative flexibility. Lawson (2006) and Cross (2023) emphasize that decisions made in the initial stages of design critically shape all subsequent steps, and that the ability to rapidly generate alternatives substantially enriches design thinking. At this stage, designers must simultaneously evaluate a wide range of considerations, including formal schema development, spatial organization, scale and proportion relationships, material potentials, and environmental performance (Lawson, 2006; Cross, 2023).

In recent years, parametric design, optimization algorithms, and various computational tools have offered designers powerful capabilities such as rapid variation generation, multi-criteria decision making (MCDM), data-driven evaluation, and performance-based comparison. While these tools are particularly effective in domains such as geometric manipulation, structural configuration, and energy performance, they remain limited in producing the more abstract design components required in the early phases, such as conceptual atmosphere, stylistic direction, material language, spatial mood, and user experience. Moreover, many of these tools rely on predefined parametric relationships, preventing them from directly visualizing concepts that designers express intuitively or verbally.

In this context, generative artificial intelligence models capable of producing visual variations from textual inputs have the potential to fill this gap left by traditional computational tools. Text-to-image models support early-stage decision-making by rapidly visualizing verbally articulated design concepts, thereby narrowing the gap between abstract design thinking and visual representation and introducing a new paradigm for design decision support.

2.4 Approaches to Sustainable Tiny House Design

The literature on sustainable tiny house design is shaped around key sustainability principles such as compact spatial organization, energy efficiency, the use of natural materials, low carbon footprint, and multifunctional spatial configurations. Studies on the Tiny House Movement have examined how minimalist living philosophies influence design decisions, demonstrating that factors such as reducing unnecessary volume, incorporating multifunctional furniture, optimizing solar exposure and natural ventilation, and minimizing environmental footprint constitute the core of the design process (Shearer & Burton, 2019). This approach extends beyond mere physical downsizing to encompass reduced resource consumption, flexible patterns of use, and the integration of user behavior into spatial configuration.

Performance-oriented design literature highlights that, in small dwellings, environmental strategies, such as strategic placement of openings, minimizing heat loss, integrating renewable energy systems, and optimizing daylight conditions must be established in the early design stages (Ford & Gomez-Lanier, 2017). In structures with limited volume, even small design decisions, such as window-to-wall ratios, façade orientation, or roof geometry, have measurable impacts on spatial comfort, energy consumption, and overall quality of life.

Therefore, early-stage decision support in tiny house design requires a framework that not only proposes spatial arrangements but also enables designers to rapidly evaluate multi-criteria options aligned with sustainability principles. Generative artificial intelligence models are emerging as complementary tools in this domain, offering the capacity to generate multiple typological alternatives within constrained space, produce material and atmospheric variations, and facilitate the visualization of environmental strategies, thus contributing a new layer of decision-support capabilities to existing design methods and approaches.

3. Methodology

This study employs a multi-stage methodology to examine the usability of text-to-image generative AI models as early-stage decision-support tools in sustainable tiny house design. The method consists of establishing controlled prompt categories, selecting the models, standardizing the generation workflow, conducting an evaluation panel, and performing a multi-criteria analysis.

The research process is structured according to the workflow illustrated in Figure 1. This workflow clearly illustrates how T2I models are integrated into the design decision-making process and strengthens the replicability of the research.

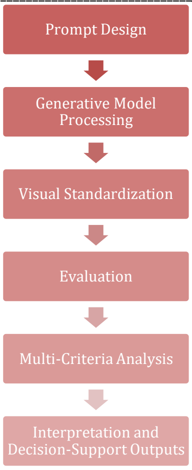


Figure 1. Workflow of the Study

3.1 Establishing Prompt Categories

All visual generation inputs used in the study were systematically defined under three main categories to ensure reproducibility:

(1) Morphological Decisions
- Plan schema (L-plan, linear plan, compact plan)
- Massing and roof form
- Opening configuration (window/door placement)

(2) Material Decisions
- Timber (CLT, solid wood), recycled composites, hybrid systems
- Emphasis on natural materials
- Surface texture and color palette

(3) Spatial Atmosphere
- Direction and intensity of daylight
- Interior illumination level
- Atmosphere: warm, neutral, or dramatic

A total of 12 controlled prompts were created for each category (Table 1), resulting in a balanced dataset suitable for cross-model comparison. This prompt set is systematic, reproducible, highly appropriate for model benchmarking, and aligned with a decision-support–oriented evaluation framework.

Table 1
Controlled Prompt Set for Sustainable Tiny House Design

Category	Prompt No	Controlled Prompt
Morphological Decisions	M1	*"Sustainable tiny house, compact single-floor plan, 24 m², south-facing façade, optimized window placement, minimal circulation area, axonometric view."*

Category	Prompt No	Controlled Prompt
	M2	*"Tiny house conceptual form, pitched roof, passive solar design, rectangular footprint, efficient space zoning, architectural diagram style."*
	M3	*"Ultra-compact tiny home, split-level interior, loft sleeping area, movable partitions, modular spatial layout, early-stage sketch aesthetics."*
	M4	*"Sustainable micro house, cube-like massing, cross-ventilation openings, shading overhangs, north-light optimization, clean linework."*
Material Decisions	MA1	*"Tiny house exterior in natural timber cladding, FSC-certified wood, exposed joinery, eco-friendly finishes, Scandinavian sustainable material palette."*
	MA2	*"Sustainable micro dwelling with recycled composite panels, low-carbon façade materials, matte textures, modern minimal aesthetic."*
	MA3	*"Tiny house interior with reclaimed wood flooring, lime-based plaster walls, breathable natural materials, warm tactile textures."*
	MA4	*"Hybrid construction tiny home: timber frame + lightweight steel joints, high-performance insulation, low-embodied-carbon material assembly."*
Spatial Atmosphere	A1	*"Interior atmosphere of a tiny house with soft daylight, diffuse north light, muted color palette, calm thermal sense, minimal furnishings."*
	A2	*"Warm and cozy tiny home interior, golden-hour sunlight, high contrast shadows, natural textures, emotionally comforting ambiance."*
	A3	*"Bright tiny house interior, clear daylight penetration, high reflectance surfaces, balanced luminance ratios, visually accessible layout."*
	A4	*"Evening tiny house interior, low-energy LED ambient lighting, warm dimmable tones, intimate spatial atmosphere, eco-friendly lighting scheme."*

3.2 Model Selection

Three widely used and technically robust diffusion-based text-to-image models were employed in this study:

- ChatGPT (DALL·E 3, OpenAI): DALL·E 3 provides exceptionally high sensitivity to textual prompts and is capable of translating them into highly detailed visual outputs (Pooja et.al.) Its ability to convert explicit design intentions directly into visual proposals makes it a powerful candidate for early-stage decision-support in architectural design.
- Copilot Designer (Microsoft): This tool offers rapid generation and stylistic diversity during conceptual design, significantly shortening the visualization process. According to Microsoft, it enables users to create and refine visual content that matches their "imagined" scene within seconds (URL-1).

- Gemini AI (Google): As an advanced generative AI platform with a multimodal architecture, Gemini AI can process visual and textual data simultaneously. Features such as high-accuracy scene interpretation, strong contextual consistency, multi-stage prompt processing, and fine-grained compositional control make it a valuable tool for architectural visualization and design research (Belaroussi, 2025).

The models were compared in terms of generation quality, prompt fidelity, capacity for producing variations, and their responsiveness to sustainability-oriented design inputs.

3.3 Visual Generation Process

The visual generation process was carried out in five stages:

- Prompt Construction: The prompt set was created by forming combinations from the three main categories.
- Image Generation: Using the same prompt set, each model produced 10 images per prompt category, yielding a total of 30 images.
- Pre-Filtering: Images containing spatial/architectural logic errors, deviating from the tiny house typology, or producing incorrect content were excluded.
- Preparation for Analysis: A final visual set was assembled for the evaluation panel. For each prompt, four images generated by each model were selected for the final dataset.

The primary objective of the workflow prepared for the visual generation process (Fig2) is to ensure equal conditions across models for comparative analysis.

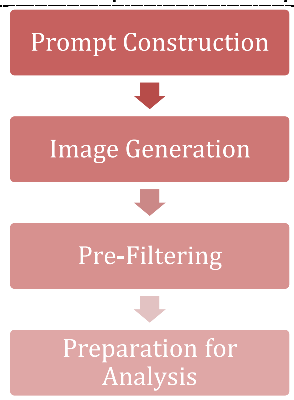


Figure 2. Workflow of the Visual Generation Process

3.1 Establishing the Evaluation Criteria

The criteria used in this study were defined with reference to the sustainable tiny house literature and fundamental architectural design principles:

- Morphological Suitability
- Performance Potential of Openings (daylighting / ventilation)
- Material / Texture Consistency
- Atmosphere / Light Quality
- Alignment with Sustainability Principles
    - Natural materials
    - Energy-efficient openings

- Compact spatial configuration
- Overall Visual and Typological Consistency

These criteria were employed both to reveal differences among models and to evaluate the decision-support potential of T2I tools in design workflows. The set of criteria was structured to enable an interdisciplinary assessment of how text-to-image models contribute to early-stage decision making in sustainable tiny house design. Morphological suitability draws on the vision-based pattern analysis literature, examining formal patterns, spatial relationships, and typological coherence (Starzyńska-Grześ et.al., 2023) Material/texture consistency corresponds to image-based assessment approaches that evaluate the accuracy and reliability of material representations in visual outputs (Chang et.al., 2025). Atmosphere/light quality is informed by key variables in computer graphics, lighting simulation, and imaging research. Sustainability alignment represents an application-oriented and multidisciplinary perspective, requiring the modeling of multi-criteria design objectives such as natural materials, energy efficiency, and compact spatial use (Elnabawi, 2025). Finally, typological consistency reflects pattern-recognition approaches that identify distinguishing features of building types, thereby assessing the overall coherence of the generated design (Pandey et.al., 2024).

4. Findings

The study presents the analysis results of the visual outputs generated using ChatGPT (DALL·E 3), Copilot Designer, and Gemini AI, with the aim of providing early-stage decision support in sustainable tiny house design. The visual sets produced based on the three prompt categories defined in the methodology—morphological decisions, material decisions, and spatial atmosphere—were evaluated in terms of architectural coherence, alignment with sustainability principles, and overall visual quality.

The findings first introduce a comparative analysis of the design patterns emerging across the prompt categories, followed by a detailed examination of each model's capacity to interpret architectural language, supported by evaluation data and observed visual characteristics.

4.1 Pre-Filtering Results

To ensure methodological transparency and preserve the validity of the evaluation panel, the raw outputs generated by the models were subjected to a pre-filtering process. At this stage, images exhibiting common generative AI issues—such as hallucination and semantic drift—were removed from the dataset.

An initial inspection of the 30 raw images produced by the three models (10 each from ChatGPT, Copilot, and Gemini) resulted in the exclusion of 18 images. A final analysis set of 12 images (4 per model) was assembled. The primary errors leading to the elimination of images were classified under three categories:

- Typological Inconsistency (Scale Error): Despite the prompt specifying a "tiny house," several models generated 2–3 storey buildings, large-footprint structures, or villa-like forms. This indicates a tendency to drift toward conventional residential typologies rather than adhering to the "compact/micro living" concept.

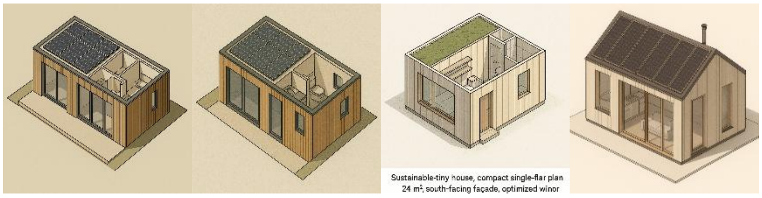
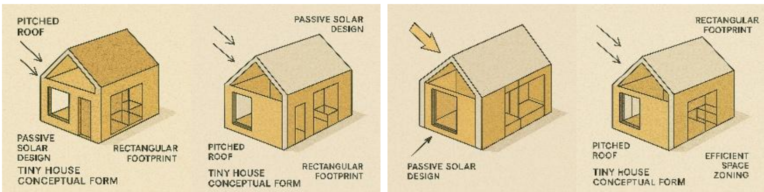
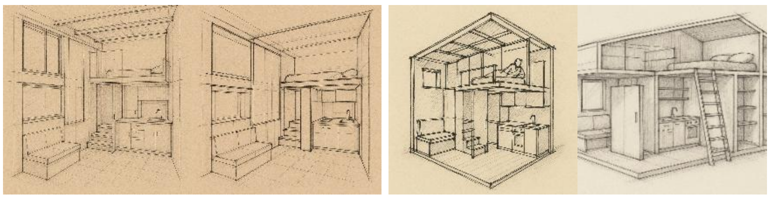
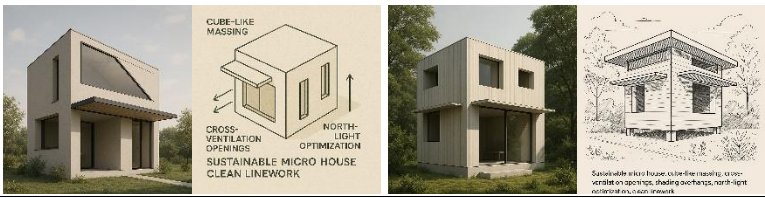
- Spatial and Architectural Logic Errors: Disconnected interior–exterior relationships, nonsensical plan organizations, improper placement of openings, physically impossible roof or mass junctions, and structural inconsistencies such as staircases terminating at walls or voids.
- Plan–Elevation Contradictions: Inconsistencies within the same image regarding scale, proportion, opening arrangement, or mass hierarchy; failure to visually reflect explicit prompt directives (e.g., "south-facing façade" or "pitched roof"); and distortions in roof geometry caused by perspective or non-Euclidean rendering artifacts.

This pre-filtering process quantitatively demonstrates that generative models still require a form of "architectural editorial supervision" and that raw outputs cannot yet be used directly as professional design substrates without critical review.

4.2 Visual Output Analysis by Prompt Categories

The visual outputs generated from the three main prompt categories defined in Table 1 (morphological, material, and atmospheric) were analyzed in detail. The resulting images (Tables 2, 3, 4) reveal how the models respond to architectural terminology and how effectively they visualize the corresponding design components within each category.

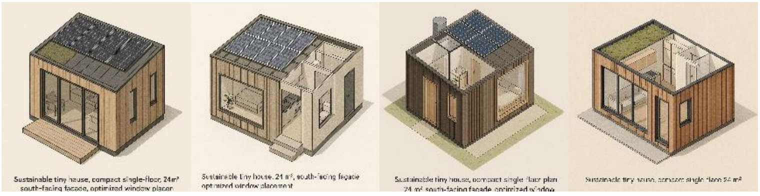
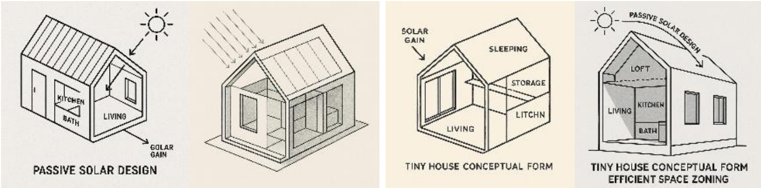
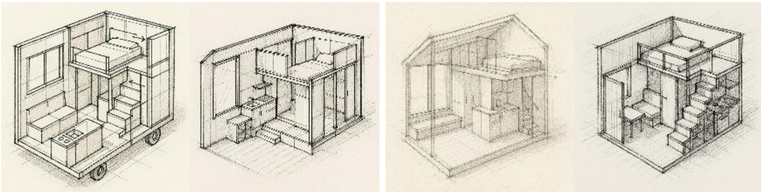
Table 2
Final Set of Outputs Generated by ChatGPT (DALL·E 3, OpenAI)

<table>
<tr><th>Prompt No</th><th>AI-Driven Visuals</th></tr>
<tr><td>M1</td><td>
</td></tr>
<tr><td>M2</td><td>
</td></tr>
<tr><td>M3</td><td></td></tr>
<tr><td>M4</td><td>
</td></tr>
</table>

Prompt No	AI-Driven Visuals
MA1	
MA2	
MA3	
MA4	
A1	
A2	
A3	
A4	

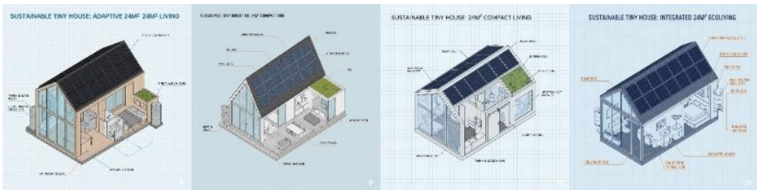
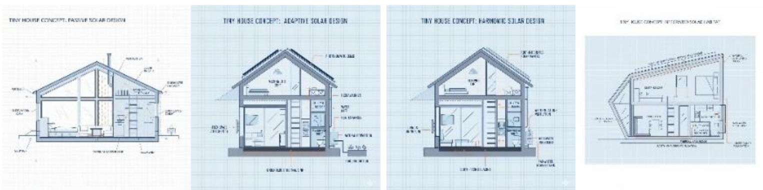
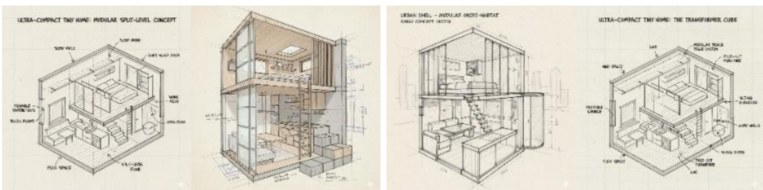
Argumentation Based Systems Journal
Volume 2, (2026) 175-192

Table 3
Final Set of Outputs Generated by Copilot Designer (Microsoft)

Prompt No	AI-Driven Visuals
M1	
M2	
M3	
M4	
MA1	
MA2	
MA3	
MA4	

Prompt No	AI-Driven Visuals
A1	
A2	
A3	
A4	

Table 4
Final Set of Outputs Generated by Gemini AI (Google)

Prompt No	AI-Driven Visuals
M1	
M2	
M3	
M4	

Prompt No	AI-Driven Visuals
MA1	
MA2	
MA3	
MA4	
A1	
A2	
A3	
A4	

The morphological prompt set defines the fundamental structural components of the dwelling, including overall form, plan schema, volumetric organization, and opening configuration. The M1–M4 prompts, which focus on morphological decisions, were designed to test structural inputs such as “compact plan,” “massing,” and “roof geometry.” The analysis indicates that text-to-image models possess a high capacity for generating massing variations. Among all categories, morphology is where differences between the models become most distinctly apparent.

- Output Patterns: ChatGPT (DALL·E 3) and Copilot predominantly visualized geometric definitions such as “L-plan” or “cube form” using perspective or isometric compositions. By contrast, Gemini AI tended to produce more technical, diagrammatic representations with a blueprint-like aesthetic, often focusing on spatial organization schemas rather than pure massing (Table 4, M1–M4). DALL·E 3 generally provided more coherent plan structures and balanced proportions, whereas Copilot Designer demonstrated strong conceptual diversity but occasionally produced spatial inconsistencies.
- Scale Problem: While massing behavior was largely successful, models frequently violated the “tiny house” scale—especially in multi-storey or loft configurations—producing dwellings that visually approached villa-scale architecture. This constitutes the most prominent source of noise within the morphological category.

The material-oriented MA1–MA4 set aimed to visualize textural distinctions such as “timber,” “recycled composite,” and “natural plaster.”

- Texture Behavior: Among all categories, material prompts yielded the most consistent responses. DALL·E 3 processed descriptors such as “reclaimed wood” or “matte texture” with high visual fidelity. However, despite its lexical accuracy, the model occasionally weakened the relationship between material representation and spatial realism in the context of small dwellings. Copilot outputs tended to present smoother, idealized (render-like) surfaces; although stylistically diverse, some examples exhibited artificiality or texture repetition. Gemini AI produced the most consistent results in this category, demonstrating high-resolution texture modeling and fine-grained detail sensitivity. Overall, the material–texture category was the second most informative in revealing technical differences between models.
- Sustainability Encoding: The phrase “sustainable material palette” was generally interpreted by the models as a combination of timber usage and vegetative elements. This suggests that generative AI relies on a visual stereotype of “green architecture” rather than a deeper technical understanding of sustainable materiality.

The atmosphere–lighting category is critical for defining the sensory qualities of interior space; therefore, it provides insight into how models simulate light behavior. The A1–A4 prompt set examined parameters such as daylight direction, color temperature, and emotional tone (cozy, dramatic, bright).

- Comparative Observation: Atmosphere generation emerged as the strongest area of performance across all models. Lighting descriptors such as “golden-hour sunlight” (A2) and “soft daylight” (A1) were the most influential parameters in enhancing spatial depth. ChatGPT (DALL·E 3) produced controlled atmospheres aligned with textual definitions of light direction and intensity. Copilot Designer also performed well—particularly in dramatic lighting scenarios—but occasionally introduced ambiguous light sources and shadow inconsistencies. Gemini AI generated the most photorealistic light and shadow distributions. While overall consistency in atmosphere–light outputs was higher than in morphological prompts, notable aesthetic differences persisted across models.

- Emotional Impact: Light–shadow dynamics often masked physical or architectural flaws (e.g., incorrect stair details), thereby increasing the perceived persuasiveness of the design.

Across categories, the highest consistency and performance were observed in the Atmosphere/Light category. The models were highly adept at transforming abstract atmospheric concepts into concrete visuals. Conversely, the greatest inconsistencies occurred in the Morphological Decisions category, where mismatches between plan and façade openings, as well as structural logic errors, were common sources of noise. The sharp contrast between Gemini AI's technical/diagrammatic approach and the pictorial approach of the other models demonstrates how prompt categories diverge depending on the model's underlying generative nature. Shared weaknesses among the three models include typological hierarchy errors, opening placement issues, and spatial logic inconsistencies in tiny house designs. The material–texture category emerged as the domain where technical performance differences were most visibly pronounced.

4.2 Evaluation Panel

At this stage of the study, the visual sets that passed the pre-filtering process were evaluated. The evaluation procedure was conducted using the six primary criteria detailed in the Methodology section, which are grounded in the literature on vision-based pattern analysis, image-based assessment, and performance simulation.

For each criterion, scores ranging from 1 (Very Weak) to 5 (Very Strong) were assigned, and the mean values were calculated. The comparative performance profiles of the models are presented in Table 5.

Table 5
Evaluation Panel

Evaluation Criterion	ChatGPT (Dall-E 3)	Copilot AI	Gemini AI
Morphological Suitability	4.2	3.8	4.5
Openings Performance	3.9	3.8	4.1
Material/Texture Consistency	4.1	4.0	4.7
Atmosphere/Light Quality	4.6	4.2	4.8
Sustainability Alignment	3.7	3.6	3.8
Typological Consistency	4.4	3.4	4.3
Overall Mean	4.2	3.8	4.4

The panel results indicate that the models exhibit notable differences particularly in morphological patterning, spatial relationships, and typological legibility. During the evaluation process, morphological suitability was assessed using the principles of vision-based pattern analysis, examining the correctness of formal patterns and spatial relations. Assessments based on Morphological Suitability and Typological Consistency revealed a clear differentiation among the models. Gemini AI received the highest score (4.5) in this category, as its outputs convey plan schemas and massing relationships in a more technical and diagrammatic manner. Its blueprint-like aesthetic makes architectural patterns more readily recognizable (pattern recognition). ChatGPT (DALL·E 3) demonstrated the most balanced performance in perspective construction and mass proportions, positioning it within a reliable range for typological coherence.

In the Material/Texture Consistency and Atmosphere/Light Quality criteria, the artistic capacity and rendering style of diffusion models were decisive, and all three models performed relatively well.

The Sustainability Alignment criterion which includes natural material use, energy-efficient openings, and compact spatial organization was the most challenging area for the models. In many cases, sustainability was interpreted superficially, often through the addition of “green façades” or “solar panels,” reflecting a form of visual greenwashing rather than a deeper understanding of environmental logic. Although some models produced visuals with a sustainable character, the underlying spatial and environmental reasoning was not always consistent.

Overall, the analysis indicates that the models do not offer a single optimal solution but instead function as specialized tools that respond to different design needs (technical diagrams vs. atmospheric representations). The evaluation results show that each model has distinct strengths and limitations across the criteria: while text-to-image generation holds strong potential for supporting early-stage design decisions, noticeable intra-model variability remains concerning morphological, typological, and sustainability-related consistency. These findings support positioning T2I models as complementary decision-support tools in early design workflows rather than autonomous design agents.

5. Conclusions

This study systematically examined the applicability of diffusion-based text-to-image (T2I) models (ChatGPT (DALL·E 3), Copilot Designer, and Gemini AI) as early-stage decision-support tools in sustainable tiny house design. The controlled prompt set structured across three primary axes, morphology, materiality, and spatial atmosphere, enabled a comparative evaluation of how these models interpret architectural inputs.

The findings indicate that generative AI models exhibit the highest level of consistency in the domain of spatial atmosphere and lighting quality. They successfully translate abstract concepts such as daylighting, color temperature, and spatial mood into visually compelling and aesthetically coherent outputs. Material and texture representation was also relatively strong; Gemini AI excelled in high-resolution surface detail, while DALL·E 3 demonstrated notable fidelity to textual prompts. In contrast, morphological consistency emerged as the weakest area. Misinterpretation of the tiny house scale, inconsistencies between plan and façade, and architectural/physical logic errors were recurrent issues observed in the generated visuals.

Despite these limitations, T2I models make substantial contributions to the early design process. By rapidly generating numerous alternatives, they expand the designer’s conceptual search space; they facilitate the early visualization of sustainability principles (natural materials, compact spatial organization, and energy-efficient openings); and they accelerate the decision-making process. In this sense, generative AI tools function not as autonomous designers but as complementary decision-support components that enrich the designer’s cognitive workflow.

Overall, the study demonstrates that diffusion-based T2I models hold transformative potential for sustainable tiny house design, yet they still require improvement in architectural coherence, typological fidelity, and environmental reasoning. As generative AI evolves toward stronger environmental inference capabilities, more accurate typological modeling, and greater structural realism, these systems are expected to become increasingly effective and knowledge-driven "intelligent design partners" in early-stage workflows. Future research should focus on multimodal pipelines integrating visual generation with performance-based tools such as energy simulations, enabling workflows that incorporate not only formal outcomes but also quantitative environmental metrics.

Author Contributions

Conceptualization, X.X.; methodology, X.X.; software, X.X.; validation, X.X.; formal analysis, X.X.; investigation, X.X.; resources, X.X.; writing—original draft preparation, X.X.; writing—review and editing, X.X.; visualization, X.X. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Conflicts of Interest

The author(s) declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The author gratefully acknowledges the support of ChatGPT for linguistic assistance in preparing this manuscript.

References

Belaroussi, R. (2025). Subjective Assessment of a Built Environment by ChatGPT, Gemini and Grok: Comparison with Architecture, Engineering and Construction Expert Perception. *Big Data and Cognitive Computing*, *9*(4), 100.

Chang, J., Lv, Y., Wang, J., Pang, H., & Liu, Y. (2025). Raster Image-Based House-Type Recognition and Three-Dimensional Reconstruction Technology. *Buildings*, *15*(7), 1178.

Chen, J., Wang, D., Shao, Z., Zhang, X., Ruan, M., Li, H., & Li, J. (2023). Using artificial intelligence to generate master-quality architectural designs from text descriptions. *Buildings*, 13(9), 2285.

Cross, N. (2023). *Design thinking: Understanding how designers think and work*. Bloomsbury Publishing.

Çelik, T. (2025). Evaluating daylight performance of AI-generated housing plans via diffusion models and climate-based simulation. *AI EDAM*, *39*, e19.

Elnabawi, M. H. (2025). Integrating Sustainability in Architectural Engineering Education: A Framework for Early-Stage SDG Alignment. *Journal of Architectural Engineering*, *31*(4), 05025004.

Ford, J., & Gomez-Lanier, L. (2017). Are tiny homes here to stay? A review of literature on the tiny house movement. *Family and Consumer Sciences Research Journal*, *45*(4), 394-405.

Gallega, R. W., & Sumi, Y. (2024). Exploring the use of generative AI for material texturing in 3D interior design spaces. *Frontiers in Computer Science*, *6*, 1493937.

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in neural information processing systems*, *33*, 6840-6851.

Lawson, B. (2006). *How designers think*. Routledge.

Pandey, S. K., Staron, M., Horkoff, J., Ochodek, M., & Durisic, D. (2024, June). Comparing Programming Language Models for Design Pattern Recognition. In *2024 IEEE 21st International Conference on Software Architecture Companion (ICSA-C)* (pp. 183-190). IEEE.

Pooja, M. M., Sulfath, P. M., & Sheena, K. M. DALL-E 3: Advanced AI image generation model. *Authorea Preprints*.

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125*, *1*(2), 3.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10684-10695).

Shearer, H., & Burton, P. (2019). Towards a typology of tiny houses. *Housing, Theory and Society*, *36*(3), 298-318.

Starzyńska-Grześ, M. B., Roussel, R., Jacoby, S., & Asadipour, A. (2023). Computer vision-based analysis of buildings and built environments: A systematic review of current approaches. *ACM Computing Surveys*, *55*(13s), 1-25.

URL-1: https://www.microsoft.com/tr-tr/microsoft-365/microsoft-designer?utm_source=chatgpt.com