ISSN 2729-7640

sciendo

ALFA
ARCHITECTURE PAPERS
OF THE FACULTY OF ARCHITECTURE AND DESIGN STU
www.alfa.stuba.sk

4/2023 (Vol. 28)
DOI: 10.2478/alfa-2023-0020

Architectural alchemy: Leveraging Artificial Intelligence for inspired design – a comprehensive study of creativity, control, and collaboration

Lenka Petráková[1*]

Vladimír Šimkovič[2]

[1,2] *Slovak University of Technology, Faculty of Architecture and Design, Institute of Public Buildings, Bratislava, Slovakia*

***Corresponding author**
E-mail: petrakova.lenka@gmail.com

Article information
Sent: Mar 18, 2023
Accepted: Nov 20, 2023

Abstract: The research paper contends that Artificial Intelligence (AI) serves as a collaborative partner in architectural design, rather than merely a utility tool. To substantiate this argument, a three-phase, nine-test investigation evaluating the strengths and limitations of two prominent AI platforms: Midjourney AI and Stable Diffusion was undertaken. These platforms synergize human creativity and AI capabilities through features like text prompts and image references, thereby fostering innovative avenues in architecture. Our analysis indicates that Midjourney AI is proficient in generating initial design concepts, largely thanks to its extensive data libraries, but is deficient in design refinement and user control. Conversely, Stable Diffusion empowers designers with greater control via features like ControlNet but sacrifices visual clarity due to its smaller generative models. Both platforms share a common flaw: an overemphasis on aesthetics and shape at the expense of functional understanding. Building upon these empirical observations, the paper outlines strategies for designers to reasonably leverage AI in optimising workflows. It confirms two key hypotheses concerning the interplay of creativity, control, and collaboration, emphasising that both human architects and AI systems benefit from iterative feedback and continuous adaptation. In summary, the study posits that AI is not just an adjunct technology but a transformative force with the capacity to fundamentally alter architectural design processes, paving the way for a new paradigm where human expertise and machine capabilities converge for enriched design outcomes.

Keywords: design, architecture, artificial intelligence, AI, creativity

INTRODUCTION

Traditional architectural design provides a human-centric and intuitive way of creating spaces based on personal creativity, experience and cultural history. The final design usually unfolds linearly with predetermined approaches set in the early stages. Sometimes, the experience can overshadow objective evaluation. In contrast, the design approach enhanced by Artificial Intelligence (AI) leverages technology to expand the boundaries of inspiration and creativity, offering new ways for exploration and innovation. AI enhances creativity by allowing architects to experiment with novel forms, structures, and ideas at an unprecedented pace. This fosters an environment where architects can explore and innovate more freely and faster without the typical constraints of manual design.

This evolution of AI began in the 1950s, and since then, it has been an ever-evolving field of research that has impacted various industries, including architecture. The term "Artificial Intelligence" (McCarthy, Minsky, Rochester, Shannon, 2006) and its meaning, the use of the human brain as a model for machine logic, were defined in 1956 at the workshop titled the Dartmouth Summer Research Project. The development of AI has faced numerous doubts from the general public and private sector over the years. However, in the 1990s and 2000s, the research gradually embraced machine learning (ML) methods, and with the rapid development of the Internet, data collection and learning capacity significantly improved. The spread of AI would not be possible without further technological advancements – GPUs (Graphic Processing Units) that allowed the operations to run in parallel rather than subsequently allowing many previously unfeasible AI projects to become a reality. The 2000s hardware change democratised the use of computation power and made AI more accessible even in user laptops, and in the 2010s, the era of deep learning emerged. The "depth "describes the increased complexity of models and the increased number of artificial neurons (Chaillou, 2022). Understanding the context of technological development and data collection is essential to appreciate that today it is the first time we, as designers, have access to such a broad library and AI-generated inspiration.

For architects, this technological leap provided exciting opportunities, such as creating entirely new designs using algorithms and machine learning based on user preferences, materiality, or experience. AI's predictive capabilities allowed for enhanced building performance and sustainability. They facilitate design-

ers in tasks such as initial concept generation, design optimisation, building performance prediction, evaluation of potential effects on energy efficiency and occupant comfort, and simulation and visualisation of the final structure (Hegazy, Saleh, 2023). An example of such objective usage in architectural design can be seen in a study of Generative conceptual design via deep learning (As, Pal, Basu, 2018), where a deep neural network (DNN) approach generated conceptual designs in architecture. The generation was based on a system evaluating and scoring designs, decomposing them into building blocks and recombining them into novel compositions.

Fig. 1. AI-generated image in version 3 (left) and version 5.2 (right). AI's understanding of floor plans is improving over time. In the latest 5.2 version, the connection of the spaces is quite logical, although it still lacks technical understanding. Text to Image generation, paid version. Prompt Floor plan of a villa. (Author: Lenka Petráková using Midjourney AI platform – paid version, 2023)

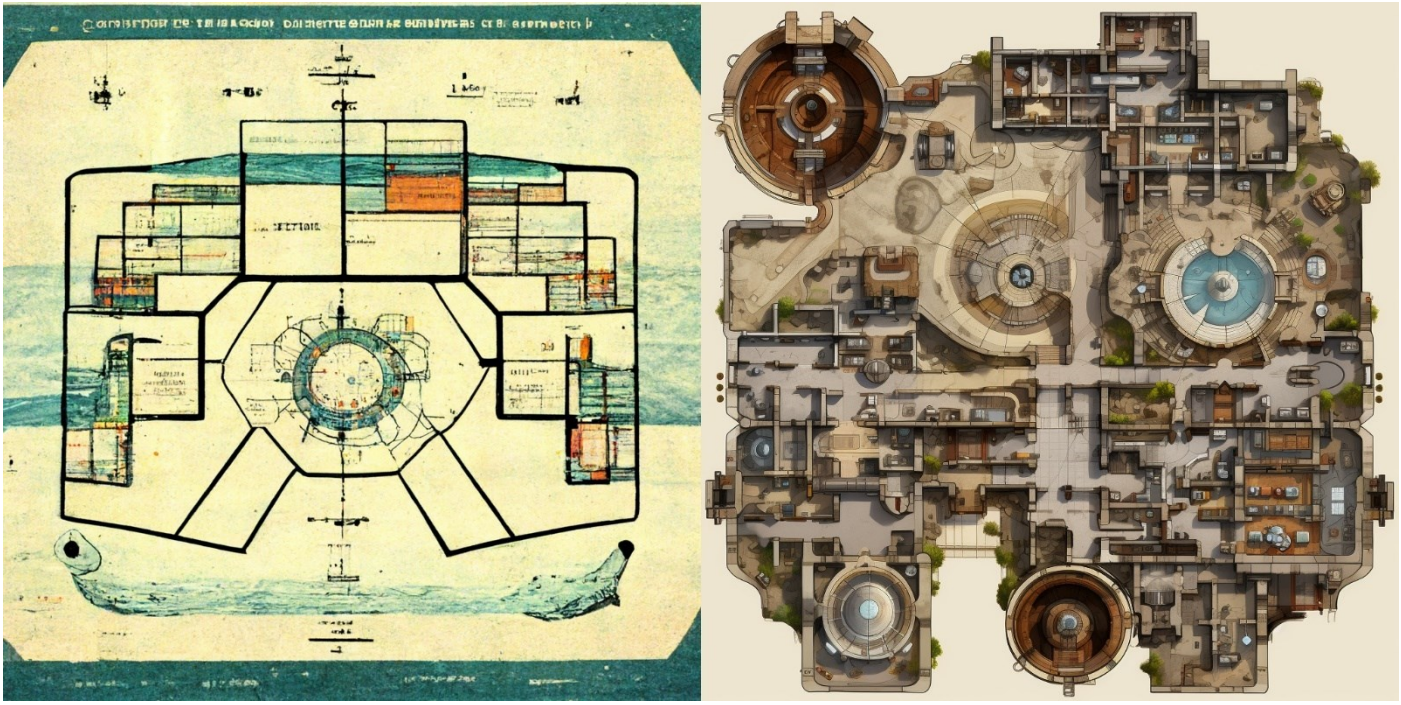

Fig. 2. AI-generated image in version 3 (left) and version 5.2 (right). We can see that the image generation was better in Fig. 1, where the language model had better historical data on spatial requirements for the villa. Text to Image generation, paid version. Prompt Floor plan of The 8th Continent, Ocean Cleaning and Research Station (Author: Lenka Petráková using Midjourney AI platform – paid version, 2023)

Within this context, two distinct AI platforms that will be examined in this paper, Midjourney and Stable Diffusion, have emerged as powerful tools in visual-driven design processes. Both platforms use AI image generation that involves previous training based on instructions and training data. This type of algorithm is called Generative Adversarial Networks (GANs). GANs help access the appearance and can be used for image generation, among others in floor plan generation (Fig. 1, 2). The GAN architecture comprises two models trained simultaneously: a generator to produce the data and a discriminator to establish the origin of the data (factual or generated) (Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville, Bengio, 2020). Despite their shared foundation, both platforms differ in the results, style, and ability to be controlled by the user. Visual-driven platforms can help in the early design stages for formal inspiration before the design is optimised and developed for its functional needs. Arguably, as the platform is taught on previous data, we could compare their proposals to designer sketches based on previous experience.

This paper delves into the effectiveness of AI-driven design approaches, exploring new ways of inspiration and innovation in the architectural sector while researching how we can control AI in the design process and use it as a tool instead of an autonomous designer. When working with the two platforms (Midjourney AI and Stable Diffusion), the questions are multifaceted and require careful consideration: How do text-to-image and image-to-image generation algorithms contribute to a more vivid visualisation of designs? How can we enable greater control and flexibility in the design process? What are their comparative strengths and limitations within the context of architectural design? How can AI's role be moderated within the design process to ensure it functions as a collaboratively interactive tool rather than an autonomous designer? By focusing on these questions, the paper aims to investigate the mechanics of these platforms, evaluating their relative capabilities and providing insights into how they can be effectively harnessed in modern design practices.

The study conducts a comparative analysis of design exercises using both Midjourney AI and Stable Diffusion. It examines the effectiveness of AI-driven design approaches in the architectural sector, ultimately transforming how architects work while recognising its limits and discussing how we could enjoy the help of AI while controlling the design process. This study adds to the expanding field of knowledge concerning the fusion of machine learning and image processing within architecture. It offers valuable insights and practical applications for design professionals and researchers in this area.

VISUAL INSPIRATION AND CREATIVITY

Human creativity is something of a mystery, not to say a paradox. One new idea may be creative, while another is merely new. What is the difference? (Boden, 2009) Human creativity can be described as our ability to develop new ideas and solutions to problems and can affect any aspect of our lives. As Boden mentioned, creativity is essential for architects, but we do not understand where it comes from. If we do not understand its origin, how can we expect AI, based on our algorithms, to introduce creativity? However, while AI may not directly produce creative solutions, it can offer diverse perspectives that might stimulate our own creativity. Essentially, it is the mental spark AI provides rather than the direct result.

AI is often defined as a tool that tries to imitate human intelligence. Margaret Boden states that AI tries to enable computers to perform operations like the human mind (Boden, 2016). It is one of the reasons why the arrival of AI and its use also brings a range of opinions on the emerging developments in artificial intelligence and machine learning and whether these technologies are beneficial or detrimental to the creative process. In architectural design, however, AI's potential to serve as an extension of human cognition or a collaborative teammate may tip the scales towards favour of enhancement rather than hindrance.

Architects often begin a design with an abstract concept and an indistinct vision of its form, forming a foundation from which a wide array of solutions can emerge (Castro Pena, Carballal, Rodríguez-Fernández, Santos, Romero, 2021). The inherent visual nature of architecture means it leans on visual cues to express ideas that might be harder to articulate verbally. Simultaneously, visual stimuli that spark creativity often inspire ideas and concepts. With technological advancements, visual inspiration extends beyond static images into interactive simulations, virtual reality, and AI-generated visuals. These tools expand the scope and depth of visual exploration, enriching the design process.

Machine learning has become increasingly popular among architects in recent years, with GANs standing out as powerful design tools. Renowned architects such as Refik Anadol and Daniel Bolojan have embraced GANs in their work, demonstrating the potential of these machine-learning frameworks that specialise in processing image-type data (Mostafavi, Tahsildoost, Zomorodian, Shahrestani, 2022). Within this technological context, AI's generation tools provide architects with a direct and innovative way to transform abstract ideas into visual representations. AI can incorporate information from various fields, including art, science, and sociology. Consequently, architectural designs are enriched with insights from cultural, ecological, technological, and social perspectives.

AI's illustrations are not confined to realism. They can act as creative symbols, encapsulating a particular sentiment or idea a space should invoke (Berg, 2022). These multifaceted visual instruments pave the way for more insightful and inventive architectural design, bridging historical and contemporary insights. Moreover, AI-generated visualisations assist in decision-making, enabling architects to harness AI as an extension of imagination rather than merely a tool for realistic rendering (PA Next Team, 2022). Balancing AI's capabilities remains challenging, especially in generating images that align with specific requirements. There are too many interpretations when the text description (prompt) lacks details. However, when the description is more detailed, the visuals could be too concentrated around them and lack the overall structure of the image, or it creates similar outcomes without available variations. This was also described by Yousif and Bolojan as prompting challenges with domain and context specificity. Despite acceptable visual outcomes, their output variety sometimes suffered, leading to similar results. Due to AI models pre-training on large datasets, not domain-specific ones (Yousif, Bolojan, 2021).

The abstract use in conceptual design often faces issues like over-concentration on details or lack of variations, which was critical primarily in the first versions of the platforms. In the early language models, the libraries used for “training” the AI were so limited that the language of all designs, no matter the input preferences, resulted in repetitive patterns. (Fig. 3) Design strategies are routed in creating new and learning from previous iterations. Therefore, unlike in other disciplines, getting the same or similar outcome twice is a failure. Designers must rigorously oversee the selection process for final outputs. However, AI holds promise in augmenting architectural creativity. Architects can innovate and refine design techniques by integrating human ingenuity with AI capabilities. As AI's role in architecture expands, it is essential to balance its potential with

its limitations, preserving the unique human touch. This paper delves into two AI platforms, Midjourney and Stable Diffusion, to explore AI's practical implications and future in architectural design.

Fig. 3. Left: Midjourney version 1 AI-generated image of a villa in the rainforest. The design is abstract, lacking material and environmental specificity. Prompt: fluid organic futuristic biologic wooden villa, futuristic shape render in V-Ray, in the Amazon rainforest, sunset hyperrealistic scene. Right: AI-generated image in Midjourney version 1, prompt: The 8th Continent Ocean Cleaning and Research Station. There are significant similarities in the architectural language compared to the villa. (Author: Lenka Petráková using Midjourney AI platform – paid version, 2022)

MATERIALS, DATA AND METHODS

This research is situated within the broader context of investigating the role of artificial intelligence in transforming natural inspiration into architectural designs. The study conducts a comparative analysis of two AI platforms, Midjourney AI and Stable Diffusion + ControlNet. The experimental framework relies on visuals from The 8th Continent, Ocean Cleaning and Research Station, which is a crucial project of extensive research on nature's influence on design. Some introductory images diverge from the project but still focus on natural inspiration, highlighting each platform's capabilities and limitations. The following hypotheses are being investigated: 1. Text-to-image and image-to-image generation algorithms will enhance the visualization of designs by providing a vivid representation, thus aiding the creative process. 2. AI's role within the design process can be moderated and managed to function as a collaborative, interactive tool rather than an autonomous designer, preserving the critical human elements of creativity and innovation.

The exploration and testing of these expectations will contribute to a nuanced understanding of how AI can be effectively harnessed in modern architectural practices, offering a balanced perspective that recognises both the promising capabilities and the inherent limitations of AI-driven design approaches. To increase the understanding of the experiment, we need to introduce the two tested platforms and the functionalities we will use.

MIDJOURNEY AI INTRODUCTION

In April 2022, a company based in San Francisco established MidJourney-AI, an extension integrated within a chat server known as "Discord®" (Salkowitz, 2022). Soon, the artists, designers and even architects shifted towards experimenting with it (Radhakrishnan, 2023). The web platform is accessed through an online chat room where one can type the command "/imagine" followed by a verbal prompt describing what one wants the AI to create. The AI quickly analyses this, filters relevant data from the database, and generates four, often abstract images, bridging the gaps in the description. While there is no strict format for "/imagine", being specific enhances accuracy. It is essential to be specific, use only positive words, describe the style of the image, reference artists, and describe the camera lens, rendering engine, or image ratio for increased precision. A single word can significantly influence the design outcome. (Fig. 4). The strength of Midjourney is in creating captivating visuals from a text; the resolution, definition, and high realistic value are hard to match today by other AI platforms due to the size of data sets it operates on.

Fig. 4. AI-generated images depict varying interpretations. On the left, the image captures flower-shaped organic domes with the prompt: "flower-shaped domes." Meanwhile, the right image, prompted with "flower domes," more closely resembles greenhouses than the intended flower-shaped structures. This difference highlights the prompt's specificity. (Author: Lenka Petráková using Midjourney AI platform – paid version)

Fig. 5. On the left is an architectural massing study created by the author. In the middle is an AI-generated image for style reference of the blend function created in Midjourney AI based on the author's prompt. On the right is a blended image of massing and style reference as an inspiration for future development. (Author: Lenka Petráková using Midjourney AI platform – paid version, 2023)

Over the last year, the Midjourney AI considerably developed by introducing various new tools. Today, designers can insert or blend images as part of their workflow. This allows inputting work-in-progress 3D models and using AI as a library of future options for the design. Describing the prompt with an image reflecting the current state of architectural projects allows for better accuracy of the models and better reading of the scale and definition of the aesthetics one seeks. (Fig. 5)

STABLE DIFFUSION INTRODUCTION

Stable Diffusion, a text-to-image deep learning model launched in 2022, stands out as an open-source model that enables end users to manipulate its code. It provides various functionalities for user control within its main interface. Although the construction of input text prompts in Stable Diffusion does not vary significantly from Midjourney AI's language structure, it offers custom models trained on various files. This increased the customisation significantly and enhanced the accuracy of the visuals in a particular style or when using one's libraries, compared to Midjourney AI. At the same time, using a data set instead of multiple libraries reduces the quality of visuals in some cases, as the data set available for generation is comparatively smaller than the one Midjourney operates on.

Stable Diffusion incorporates ControlNet, which enhances user control through additional conditions and inputs based on image recognition. While ControlNet improves the precision of the generated images, it does so at the expense of flexibility since it is designed to adhere to a particular set of parameters. To utilise ControlNet, a reference image must be loaded onto the canvas, and then a control type must be used to determine what information to extract from it. Various options include Pose, Depth Mask, Line Art, Reference, Canny Edge Detector, Normal Maps, Scribble, Segmentation, and Colour Grid. Depending on the chosen control type, the reference image is processed by ControlNet to form a foundation that is subsequently altered according to the provided text description. Within each control type, the amount of information extracted can be fine-tuned by selecting the model type. With depth mask generation, different models can generate varying levels of detail from the original image (Fig. 6) or by adjusting the control weight of ControlNet concerning the text prompt. This gives the designer flexibility to concentrate on a desired option.

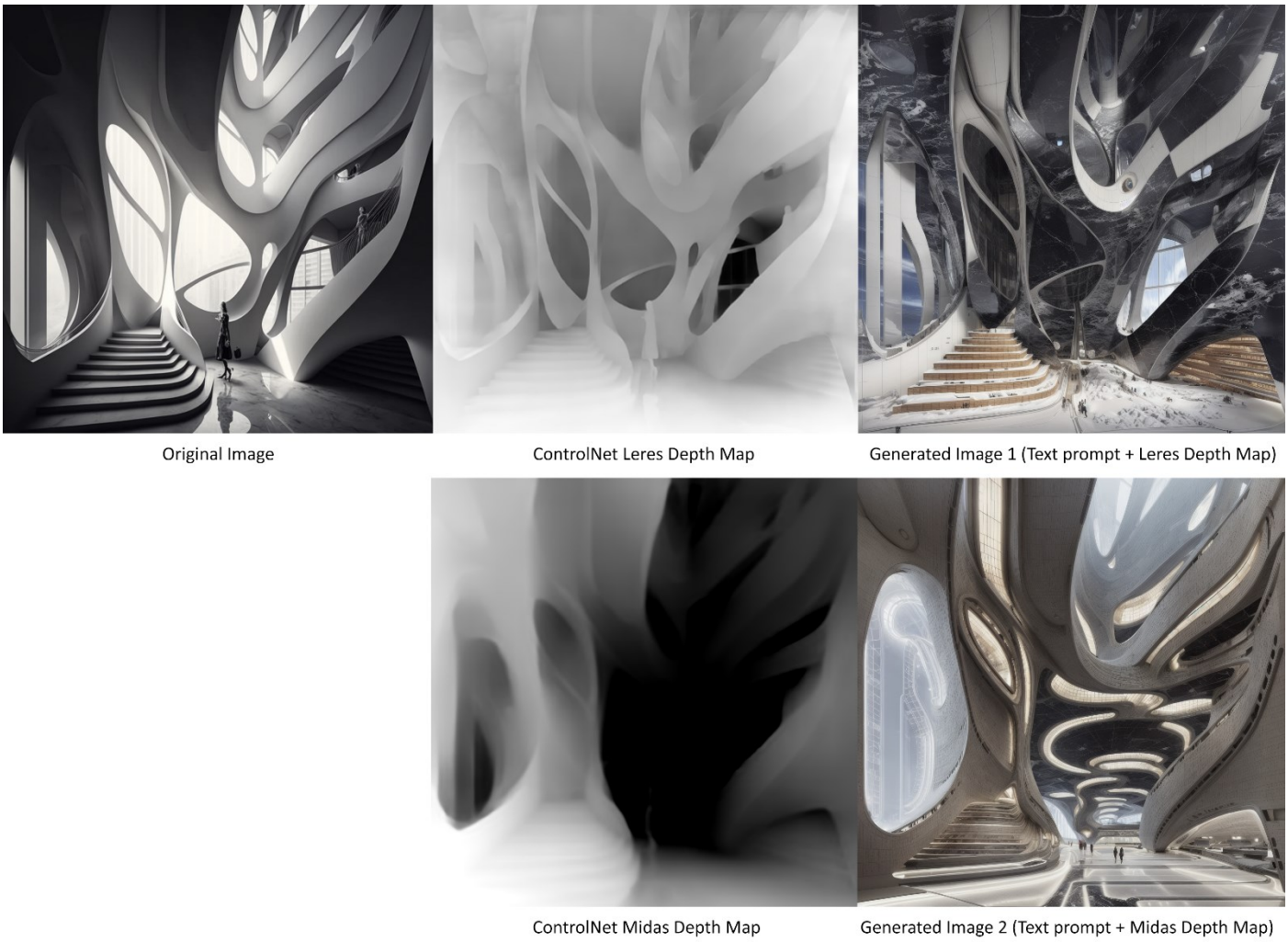

Fig. 6. Test with Depth Control Type in ControlNet. Two depth masks were generated from the original image – Leres (detailed depth map) and Midas (limited depth map to foreground). We can see that the less information extracted from the original image, the more freedom is given to AI to suggest possible designs. (Author: Lenka Petráková using Stable Diffusion Platform a111 for PC, 2023, CC0 1.0 Universal Public Domain Dedication)

EXPERIMENT

The design experiment integrates visuals of interiors with elements inspired by nature, exploring whether AI has the potential to expedite the design process by harnessing and adapting this inspiration within predefined geometries. To assess the strengths and weaknesses of this approach, three distinct natural inspirations were evaluated: **Environmental**: incorporating green spaces into the design. **Formal**: introducing shapes inspired by nature into the design. **Material**: integrating the material quality of wood into the design. The experiment is conducted on two distinct AI platforms, utilising the same input data – an image of an interior (Fig. 7) and corresponding descriptions of the references. By reason of differences in the workflows of the two platforms, the steps in the experiment vary between them.

ENHANCED INSPIRATION BY MIDJOURNEY AI

In Midjourney AI, we must employ the blend function to use a reference image as an input. This particular function prohibits the addition of further text, accepting only images. To encapsulate the natural inspiration through this function, we need to initially generate three distinct text-to-image outputs that carry the visual essence of the inspiration. However, the blend function does not permit us to control the influence of each reference on the final result.

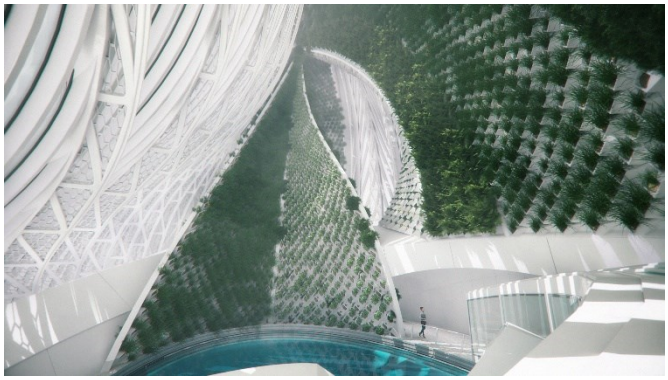

Fig. 7. Interior Render of the Greenhouses in the project The 8th Continent, Ocean Cleaning and Research Station. (Author: Lenka Petráková using Midjourney AI platform – paid version, 2023)

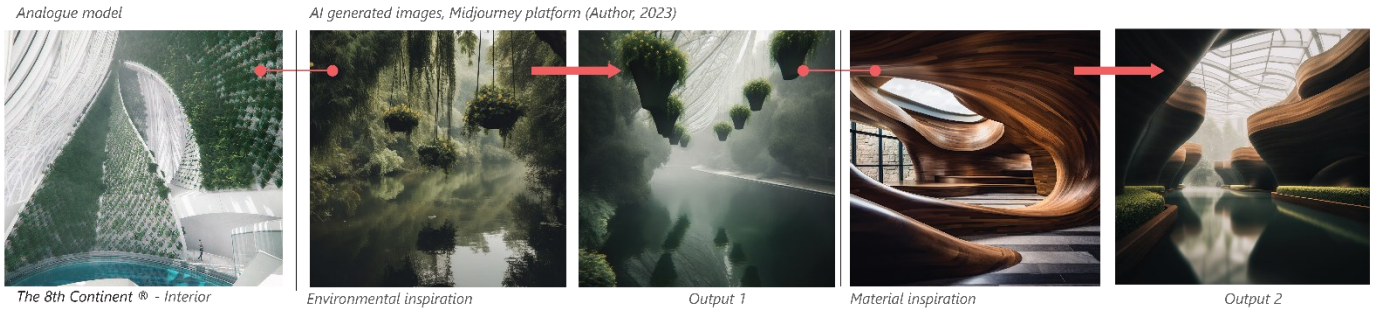

Fig. 8. Steps of Test 1 of environmental inspiration incorporated in interior design by blending the designed interior with AI references. (Author: Lenka Petráková using Midjourney AI platform – paid version, 2023)

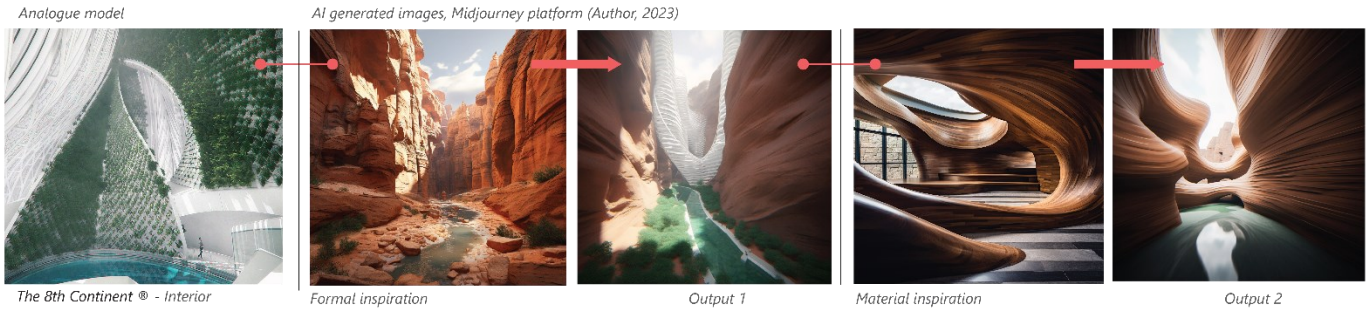

Fig. 9. Steps of Test 2 of formal inspiration incorporated in interior design by blending the designed interior with AI references. (Author: Lenka Petráková using Midjourney AI platform – paid version, 2023)

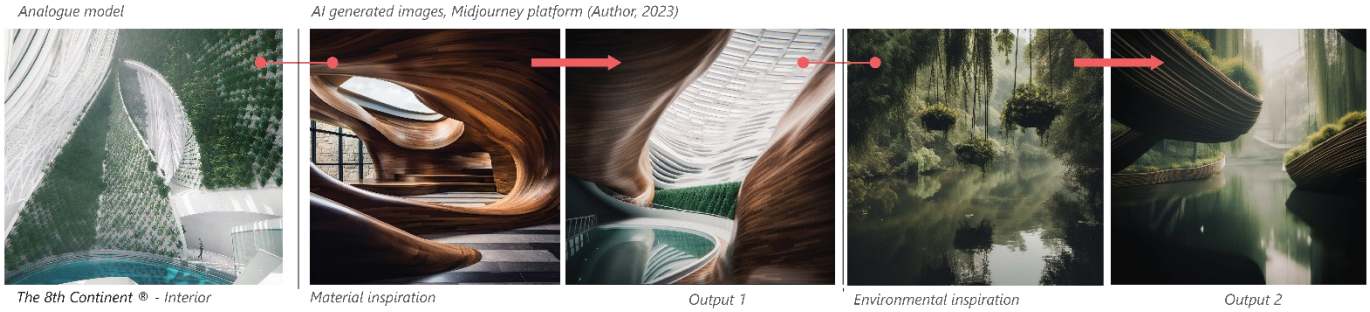

Fig. 10. Steps of Test 3 of material inspiration incorporated in interior design by blending the designed interior with AI references. (Lenka Petráková using Midjourney AI platform – paid version, 2023)

In **Test 1 – Environmental Inspiration** (Fig. 8), an environmental reference image was created based on the prompt "lake between hanging plants," reflecting the intention for the greenhouse to incorporate a pool and green walls in the interior. The initial result of merging the interior and environmental references predominantly reflects the environmental aspect, with only the facade elements representing the interior. The resulting image from the initial test run does not incorporate the material or formal aspects of the interior. To address this, a second generation was executed, wherein the first output was blended with another natural inspiration reference image – in this instance, a material reference. The second output reveals more detail, such as the striation of the wooden elements and the pool appearing more man-made than organic. However, the image lacks an understanding of the original scale of the interior.

In **Test 2 – Formal Inspiration** (Fig. 9), a formal reference image was created based on the prompt "canyon with smooth stone walls," This reflected the desire for the greenhouse to embody a language of fluid shapes in the interior. A second blend was executed to better integrate architectural elements into the final result. Here, the first output was blended with another natural inspiration reference image, specifically a material reference. However, the second output did not significantly improve. The composition remained heavily influenced by natural resemblance and failed to incorporate sufficient architectural qualities from the interior input or the material reference.

In **Test 3 – Material Inspiration** (Fig. 10), a material reference image was crafted based on the prompt "modern interior photography, wooden details." This depicted the intention for the greenhouse to include wood elements significantly since the original proposal did not demonstrate any specific materiality. The first output skilfully integrated both references, forming a feasible alternative for the interior with elements from both references present. However, the image lacked natural references, such as planting. To address this, a second test was conducted, incorporating the natural reference that had been previously generated. Unfortunately, the natural imagery supplanted the architectural qualities in the second output.

ENHANCED INSPIRATION BY STABLE DIFFUSION

To test the abilities of Stable Diffusion accelerating the design process by offering inspiration, we will use a combination of text prompts with the ControlNet depth map and segmentation control types to allow us to reference the image of a Greenhouse (Fig. 7). Stable Diffusion utilises the interior image (Fig. 7) solely for depth map or segmentation information in its image recognition process, unlike Midjourney, which can incorporate elements from the existing image and blend them with the text description. As a result, it becomes necessary to define some of the architectural elements in the Prompt to avoid losing all the architectural features. As previously discussed, Stable Diffusion operates using models ("checkpoints"). These checkpoints are pre-trained weights tailored for generating either general images or those of a specific genre. For this experiment, we will utilise two models: "Architecturerealmix," trained on architectural databases, and "Deliberate," designed for a general style.

DEPTH MAP TESTS

Test 4 – Environmental inspiration. (Fig. 11) In this experiment, two models were evaluated: "Architecturerealmix" (Fig. 11, first row) and "Deliberate" (Fig. 11, second row). To prepare the depth map for the subsequent test, we trialled two preprocessors: "Midas" (Fig. 11, third row, two images on the right) and "Leres++" (Fig. 11, third row, two images on the left). The Prompt "Green walls, mixed plants, Hydroponic planting, lush green, artificial lake, greenhouse, glass facade" was used to align environmental inspiration with architectural components. The ControlNet's control weight was adjusted to 0.5 and 1 for each generated output. This was done to analyse how the AI interprets interior details from the source image. Based on the initial test results, it was observed that "Midas" allowed the AI a relatively greater degree of freedom at both control weights, 1 and 0.5. The "Architecturerealmix" model incorporated shapes and details that were comparatively more relevant and intriguing for our study. Therefore, we focused exclusively on the "Midas" preprocessor and the "Architecturerealmix" model in subsequent tests.

Test 5 – Formal Inspiration. (Fig. 12) Formal inspiration was guided by the prompt "Stone details, canyon geometry, stone canyon, Hydroponic planting, lush green, man-made lake, greenhouse, glass facade." It is observable that although the geometric massing aligns with the original interior thanks to the depth mask, there is not much innovation in the geometry itself, with changes mainly in material replacement. To explore further possibilities, the control weight of the depth mask was reduced from 1 to 0.5. This adjustment allowed the canyon-like structures to become more dominant, leading to formal inspiration produced by the AI.

Test 6 – Material Inspiration. (Fig. 13) Material Reference utilised the prompt "Wood details, modern interior, Hydroponic planting, wooden, lush green, artificial lake, greenhouse, glass facade." The geometry was consistently maintained throughout the generations. However, the materials and mainly the greenery were applied randomly, as the only references derived from the image pertained to depth and geometric massing. In a subsequent step, the control weight of the depth mask was reduced from 1 to 0.5. Although some interior elements were protected, the image changed the character from interior to exterior, reducing the relevance of the interior design inspiration.

SEGMENTATION TESTS

The prompts on tests 7, 8 and 9 are identical to the corresponding tests 4, 5 and 6. The difference is that the referenced image is being switched from Depth Mask to segmentation control type in the ControlNet. In all tests, the Ofade20k preprocessor defines the colour map for segmentation in the first step. The subsequent three tests are carried out on custom-made colour maps, with the colour mapping following the prescribed uses for segmentation. Only the categories that AI initially recognised in the image were used for this experiment. R 230, G 230, B 230 (Bright Grey) – Windows; R 204, G 255, B 4 (Green) – Plants; R 120, G 120, B 120 (Dark Grey) – Wall; R 61, G 230, B 250 (Blue) – Water; R 80, G 50, B 50 (Purple) – Flooring. From the colour definition, we can see that the data set recognises objects rather than individual materials. The materiality can be defined further by the Prompt.

Test 7 – Environmental inspiration (Fig. 14). The segmentation map generated by the Ofade20k preprocessor from the interior image (Fig. 7) is relatively simple, categorising the image into three main elements: wall, plants, and water. However, the generated image in the first step captures the primary material and geometric elements but lacks detail. Therefore, the segmentation mask was manually edited in Adobe Photoshop to refine the result. Initially, we separated all windows from the wall category and kept only the solid areas around the pool

under the wall category. Subsequent edits added finer details to the glass facade to align it with the initial design. Among all results, the environmental inspiration was correctly managed, and the planting references are of the correct type and scale and applied on the correct surfaces.

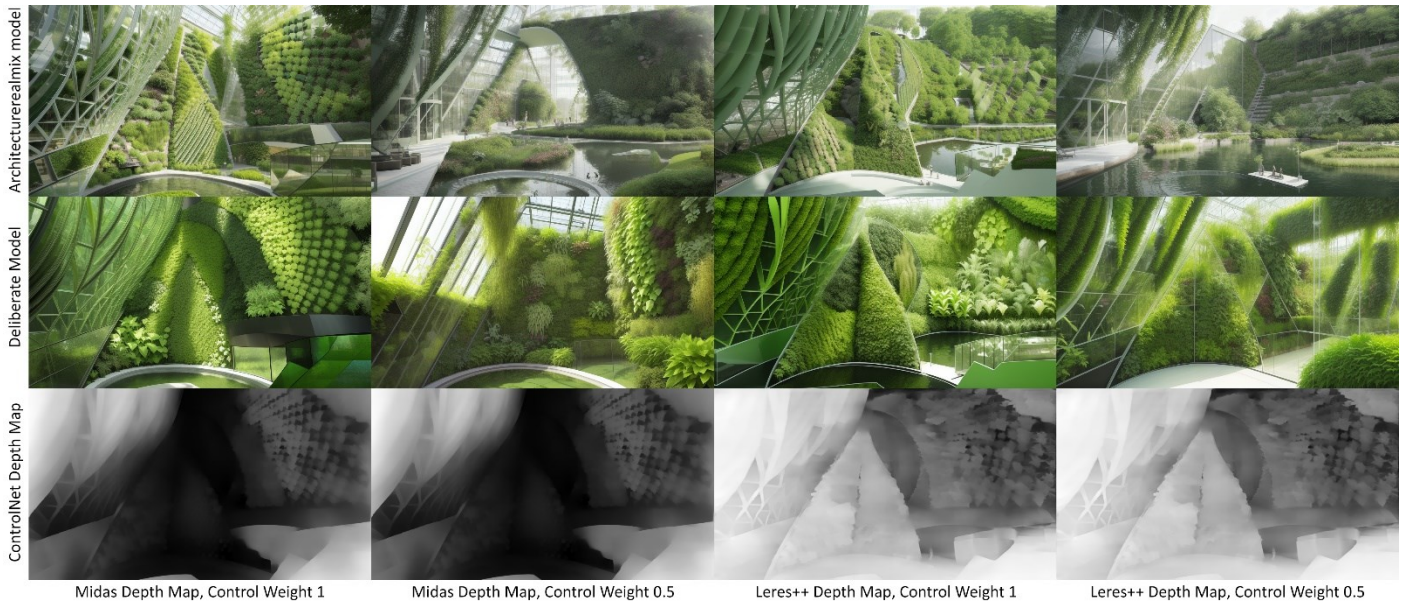

Fig. 11. Steps of Test 4, Environmental inspiration. Incorporating blending the interior design reference as a depth mask with a text description of the environmental inspiration. (Author: Lenka Petráková using Stable Diffusion Platform a111 for PC, 2023, CC0 1.0 Universal Public Domain Dedication)

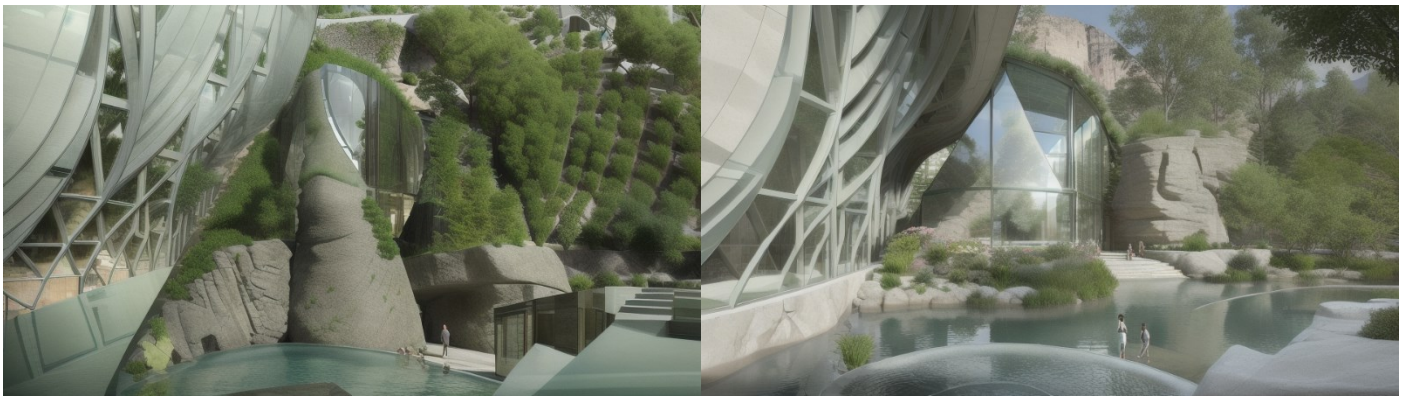

Fig. 12. Steps of Test 5, Formal inspiration. Incorporating blending the interior design reference as a depth mask with a text description of the formal inspiration. Midas preprocessor, control weight 1 (left) and control weight 0.5 (right). (Author: Lenka Petráková using Stable Diffusion Platform a111 for PC, 2023, CC0 1.0 Universal Public Domain Dedication)

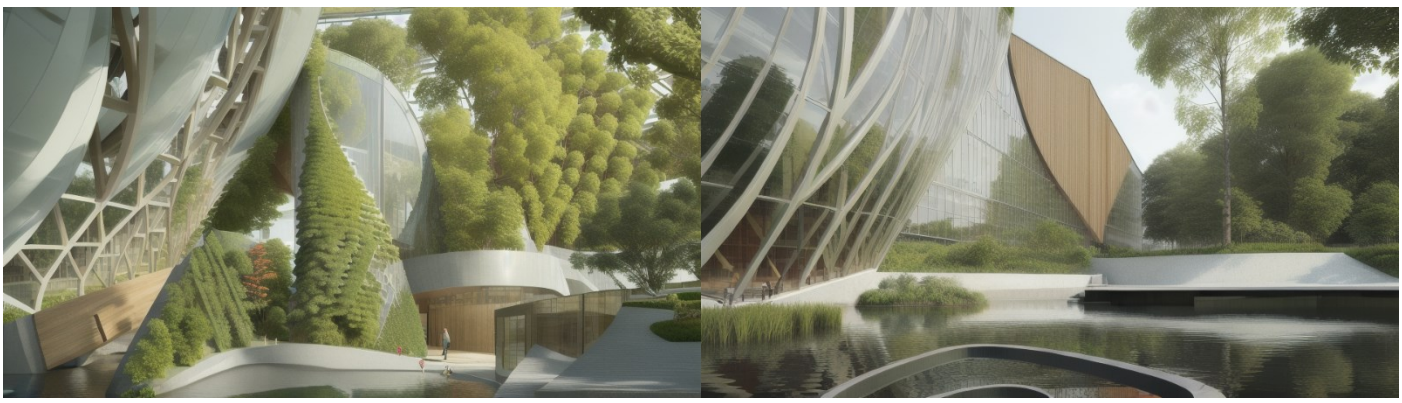

Fig. 13. Steps of Test 6, Material inspiration. Incorporating blending the interior design reference as a depth mask with a text description of the material inspiration. Midas preprocessor, control weight 1 (left) and control weight 0.5 (right). (Author: Lenka Petráková using Stable Diffusion Platform a111 for PC, 2023, CC0 1.0 Universal Public Domain Dedication)

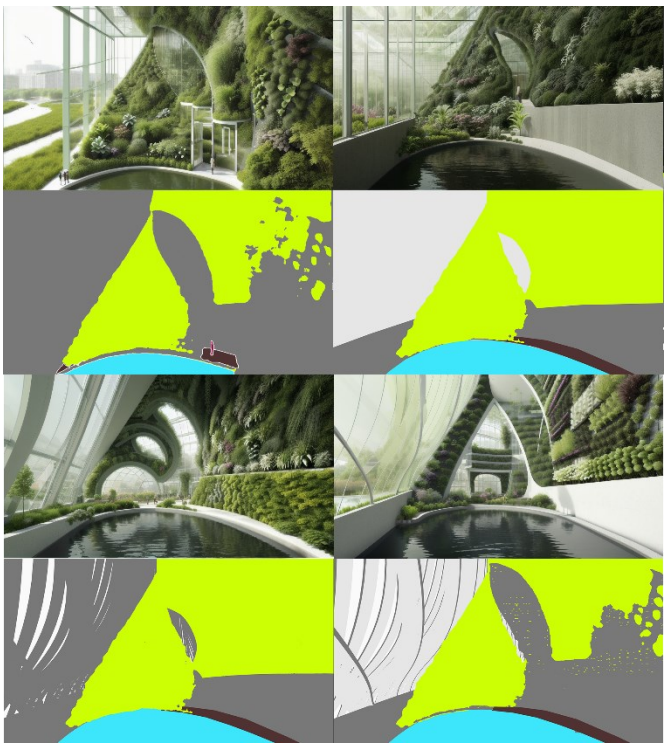

Fig. 14. Steps of Test 7, Environmental inspiration. Incorporating blending the interior design reference as a colour map defining the geometry types with a text description of the environmental inspiration. (Author: Lenka Petráková using Stable Diffusion Platform a111 for PC, 2023, CC0 1.0 Universal Public Domain Dedication)

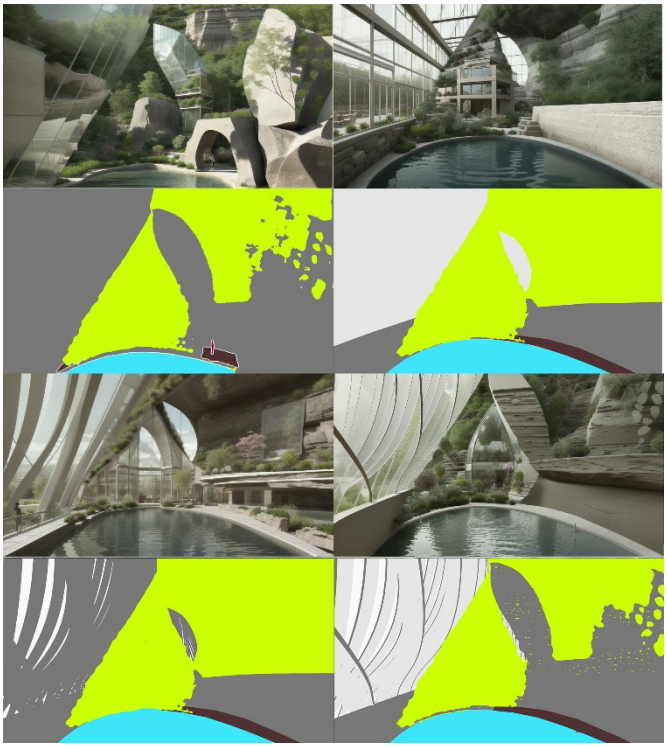

Fig. 15. Steps of Test 8, Formal inspiration. Incorporating blending the interior design reference as a colour map defining the geometry types with a text description of the formal inspiration. (Author: Lenka Petráková using Stable Diffusion Platform a111 for PC, 2023, CC0 1.0 Universal Public Domain Dedication)

Test 8 – Formal Inspiration (Fig. 15). The same segmentation masks were used as in test 7. We can observe a good understanding of the water and facade references in all the results. However, concerning the formal inspiration, the material is an applier on flat surfaces and lacks 3D definition.

Test 9 – Material Inspiration (Fig. 16) employed the same segmentation masks as Test 7. Interestingly, when wood was specified as the interior material, the AI replaced the original plant types with forest foliage and ignored the Prompt description of the hydroponic planting. Despite this change, the AI still adhered to the basic geometric outlines from the segmentation map and applied the materials accurately following the given descriptions.

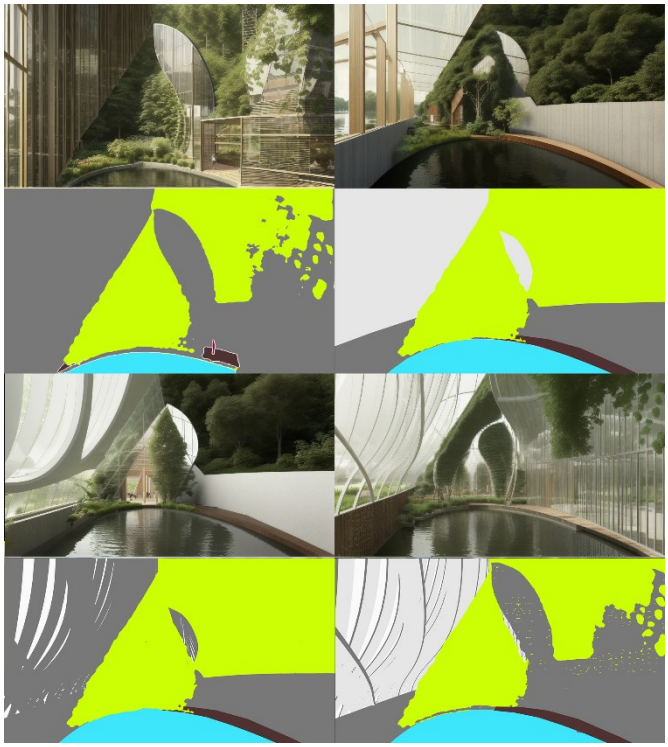

Fig. 16. Steps of Test 9, Material inspiration. Incorporating blending the interior design reference as a colour map defining the geometry types with a text description of the material inspiration. (Author: Lenka Petráková using Stable Diffusion Platform a111 for PC, 2023, CC0 1.0 Universal Public Domain Dedication)

RESULTS

Nine tests were conducted to establish the effectiveness of AI tools in architectural design and designers' ability to control the AI to achieve bespoke solutions for particular conditions. The aim was to investigate the possibilities and limitations of referencing natural inspirations within a predefined interior context. The nine tests were categorised into three distinct groups. In tests 1 to 3, where the Midjourney AI platform was employed. The process commenced with generating natural inspiration references through text Prompts. Tests 1-3 used Midjourney AI to generate images from text prompts. Tests 1 and 2 used simple prompts without architectural context, leading to perspective and image quality issues. Test 3 specified wood material for the interior, resulting in better proportionality in Output 1. A second blending in Tests 1 and 2 improved details, while in Test 3, a second blend using an environmental reference resulted in Output 2, which augmented the architectural qualities in the given example.

Several conclusions can be drawn from executed tests. While the platform excels in generating vivid imagery from text descriptions, the blending operations tend to diminish resolution. Moreover, it is not possible to adjust the weight between the references. Despite diverging from the original references, the

images met the intended objectives. The Midjourney AI suggested formal, material, and environmental modifications to the provided interior sample. Although the inspiration could not be fully utilised, elements of it could be subsequently integrated, considering the appropriate scale and context. The platform offers benefits in idea generation but is deficient in allowing the designer to control or apply the outcomes precisely. This underscores the platform's value in conceptual development while highlighting areas for potential refinement in implementation.

In Tests 4-6, we utilised the Stable Diffusion platform, incorporating ControlNet, depth map controls, and various models. This setup permitted both image and text inputs with the flexibility to adjust their respective weights. Initially, a depth mask control weight of 1 resulted in rigid and inaccurate outputs, revealing the AI's tendency to follow shapes strictly. To counter this, we standardised the control weight to 0.5 across all tests, which improved the creative output. Our findings showed that Stable Diffusion offered greater control compared to Midjourney AI. Using models trained on architectural libraries made the output more relevant and valuable for interior design. While we had the flexibility to adjust input weights, the AI still misunderstood the function of the space, focusing mainly on shapes. Despite this, the geometry, scale, and the text prompt, were more accurately captured, making Stable Diffusion better suited for case-based inspiration. However, additional designer involvement is needed to filter the results for the final application.

In tests 7–9, we used the Stable Diffusion platform along with ControlNet's segmentation control type. This setup allows specific functions to be applied to different image parts based on a designated colour map. Using a generic colour map, our first test resulted in poor comprehension of the reference spaces. To improve this, we manually refined the colour map in later tests, enhancing the depiction of interiors and material application. Achieving a balance between detail and AI's creative freedom is crucial. Over-specifying details could make the AI's unique contributions negligible, diminishing its value for inspiration. In our tests, we maintained a level of abstraction in the colour maps, which led to innovative suggestions for interior development regarding materials, form, and environmental aspects. While the generated images showed high applicability, further designer oversight is required to assess the feasibility of the proposed shapes for future design stages.

In summary, AI tools offer promising architectural design possibilities, serving as inspiration generators and collaborators that can extend human capability. However, they are not without limitations – most notably, there is the inability to fully comprehend functional spaces and offer designers complete control. Despite these constraints, both platforms have unique strengths and can significantly contribute to different design development phases.

DISCUSSION

Regarding contributing to a more vivid visualisation of designs, the text-to-image and image-to-image generation algorithms can generate robust imagery from text descriptions. However, the quality may vary with the weight and control of references. To enable greater control and flexibility, we found that adjusting the control weight of depth masks or using the ControlNet segmentation control type can offer different levels of influence over geometry, materials, and overall design. Comparatively, Midjourney AI demonstrates strength in generating ideas but may lack specific application control, while Stable Diffusion with ControlNet segmentation enables more targeted control over functions within the referenced image. The comparative study of these platforms highlights the potential of AI as a collaborative, interactive tool in architectural design rather than an autonomous designer. It provides insight into the balance between creative freedom and control, aligning AI's role with specific design intentions and encouraging the exploration of unfamiliar design patterns.

A novel approach could lie in combining the strengths of both platforms. This can be achieved by creating sketches from text-to-image or a blend of images in Midjourney AI and exporting these as image references within Stable Diffusion (Fig. 17). By doing this, designers may incorporate more realistic materials and details. This approach can enhance geometry readability while preserving the creative flair of Midjourney AI. Various control types, such as depth masks, allow designers to define the desired information level and control, balancing inspiration and precision. However, as evidenced during our design experiments, striking this balance requires careful judgment. Designers must decide how much control to cede to AI and how much alteration to accept. Although AI may not always align with human intentions, its creative freedom could be vital, spurring new design suggestions. This contrasts with conventional visualisation tools, which might only replicate what is already planned. Integrating Midjourney AI and Stable Diffusion could lead to a more balanced and enriched architectural practice. By augmenting the strengths of human intuition with AI's vast possibilities, designers can create more vivid design sketches with greater control and rapidity. This hybrid approach offers a pathway toward visualising designs, exploring unfamiliar territories, and breaking through conventional design patterns. It underscores the potential for AI to function as a collaborative, interactive tool, contributing to a dynamic interplay between creativity and control in architectural design.

CONCLUSION

Technology is evolving rapidly, and we can experience its progress on a daily basis. Soon, we could have more ways in which we benefit from the use of AI in architecture. However, we should not forget to review, analyse and continue working with the AI results to achieve the quality control and the development we seek. As McQuillian argues, there is no intelligence in artificial intelligence. Even though its technical name is machine learning, it is simply mathematical minimisation (McQuillian, 2018). This paper delved into various methodologies for integrating AI within the architectural realm, discussing its potentialities and constraints. We identified potential roles and implementations for platforms like Midjourney AI and Stable Diffusion. Each of these platforms presents distinct advantages in the domain of design visualisation. The tests revealed several nuances:

1. Stable Diffusion image-to-image algorithms offer greater control over design iterations, especially when combined with ControlNet and some control types like depth map or segmentation control types. This allows for geometry adaptations, aligning more closely with architectural intentions and providing a more practical framework for design exploration. 2. Adjusting control parameters, like the depth mask weight, can effectively bridge the gap between creative freedom and design precision. 3. A synthesis of the two platforms' capabilities suggests a promising avenue: taking initial, often avant-garde design outputs from Midjourney and refining them in Stable Diffusion for more grounded, practical, and detailed design solutions.

These tests underscore the indispensable role of human discernment, selection, and iterative adjustment in harnessing AI's capabilities. The term 'extended intelligence' resonates today, echoing the synergistic potential between AI and human cogni-

tion, where AI functions in tandem with human thought, effectively becoming an extension of it (Leach, 2021). Finding the balance between AI and human touch is essential to fully utilising the potential of AI in architecture. Design is likely the most sophisticated aspect of human intelligence (Gero, 1991). Our investigations suggest that a well-calibrated AI can amplify human creative instincts, inspiring without overshadowing the design process. Despite the challenges, tested platforms Midjourney AI and Stable Diffusion offer innovative visualisation pathways, pushing the boundaries beyond conventional human sketches.

Both research hypotheses were corroborated by our tests, which spotlighted the harmony of creativity, control, and collaboration. The proposed hybrid workflow mirrors traditional processes, emphasising goal-setting, iterative refinement through feedback loops, and oversight. Learning remains a two-way street as we navigate this digital transformation: designers adapt to platforms, and algorithms evolve through iterative feedback. In conclusion, the integration of AI, as exemplified by platforms like Midjourney and Stable Diffusion, is not merely a technological advancement but a creative catalyst that redefines the architectural process. However, creativity is a quality that cannot be automated. (Mello-Klein, 2022) By aligning computational capabilities with human intuition and creativity, we foresee a future in architecture that is as artistically profound as it is technologically advanced.

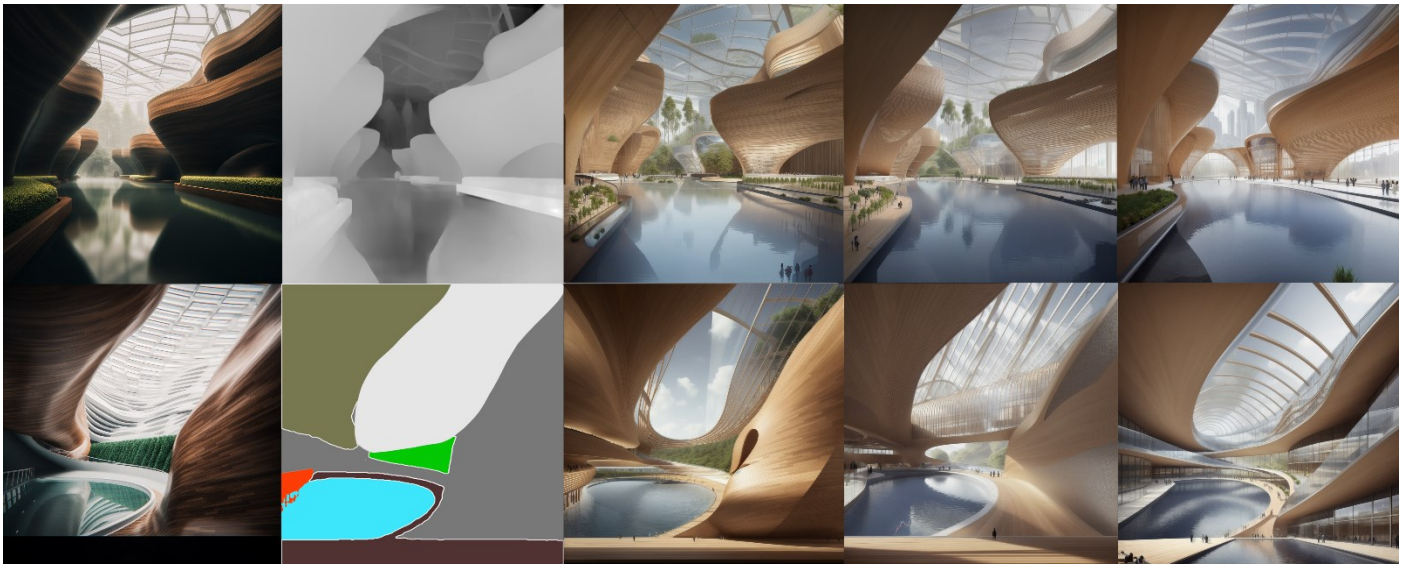

Fig. 17. Using Midjourney AI output image as input for Stable Diffusion. In the first row, use of Depth Map control type; in the bottom row, use of Segmentation control type. Both types were tested on three control weights – 2, 1 and 0.5 from left to right. (Author: Lenka Petráková using Stable Diffusion Platform a111 for PC, 2023, CC0 1.0 Universal Public Domain Dedication)

References

As, I., Pal, S., Basu, P. (2018) "Artificial intelligence in architecture: Generating conceptual design via deep learning", International Journal of Architectural Computing, 16(4), pp. 306–327. https://doi.org/10.1177/1478077118800982

Berg, N. (2022) "AI tools like DALL-E 2 and MidJourney are helping architects—and their clients—design new buildings", Fast Company. [online] Available at: https://www.fastcompany.com/90780871/ai-tools-like-dall-e-2-and-midjourney-are-helping-architects-and-their-clients-design-new-buildings [Accessed: 28 Aug 2023]

Boden, M. A. (2009) "Creativity in a nutshell", Think, 5(15), pp. 83–96. https://www.doi.org/10.1017/s147717560000230x

Boden M. A., (2016) "AI: Its Nature and Future", Oxford University Press, New York, USA.

Castro Pena, M. L., Carballal, A., Rodríguez-Fernández, N., Santos, I., Romero, J. (2021) "Artificial intelligence applied to conceptual design. A review of its use in architecture", Automation in Construction, 124, 103550. https://doi.org/10.1016/j.autcon.2021.103550

Chaillou, S. (2022) "Artificial Intelligence and Architecture, From Research to Practice", Birkhäuser Verlang, Basel, Switzerland. https://doi.org/10.1515/9783035624045-fm

Gero, J. S. (1991) "Ten problems for AI in design", presented at Workshop on AI in Design, IJCAI-91.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y. (2020) "Generative adversarial networks", Communications of the ACM, 63(11), pp. 139–144. https://www.doi.org/10.1145/3422622

Hegazy, M., Saleh, A. M. (2023) "Evolution of AI role in architectural design: between parametric exploration and machine hallucination", MSA Engineering Journal, 2(2). https://www.doi.org/10.21608/msaeng.2023.291873

Leach, N. (2021) "Architecture in the Age of Artificial Intelligence ", Bloomsbury Publishing, London, UK. https://www.doi.org/10.5040/9781350165557

McCarthy, J., Minsky, M. L., Rochester, N., Shannon, C. E. (2006) "A Proposal for Dartmouth Summer Research Project On Artificial Intelligence, August 31, 1955", AI Magazine, 27(4). https://doi.org/10.1609/aimag.v27i4.1904

Mello-Klein, C. (2022) "Artificial intelligence is here in our entertainment. What does that mean for the future of the arts?", Northeastern Global News. [online] Available at: https://news.northeastern.edu/2022/09/09/art-and-ai/ [Accessed: 28 Aug 2023]

Midjourney (2023), San Francisco, California, USA. [online] Available at: https://www.midjourney.com

Mostafavi, F., Tahsildoost, M., Zomorodian, Z. S., Shahrestani, S. S. (2022) "An interactive assessment framework for residential space layouts using pix2pix predictive model at the early-stage building design" (ahead-of-print), Smart and Sustainable Built Environment, Emerald Publishing. https://www.doi.org/10.1108/SASBE-07-2022-0152

PA Next Team (2022) "A (Mid)Journey To The Virtual World Of Hassan Ragab", Parametric Architecture. [online] Available at: https://parametric-architecture.com/a-mid-journey-to-the-virtual-world-of-hassan-ragab/ [Accessed: 28 Aug 2023]

Radhakrishnan, M. (2023) "Is Midjourney-Ai the New Anti-Hero of Architectural Imagery & Creativity?", Global Scientific Journals, 11(1). [online] Available at: https://www.globalscientificjournal.com/researchpaper/Is_MidJourney_AI_a_new_Anti_hero_of_Architectural_Imagery_and_Creativity_.pdf

Salkowitz, R. (2022) "Midjourney Founder David Holz On The Impact Of AI On Art, Imagination And The Creative Economy", Forbes Media, Business, Hollywood and Entertainment. [online] Available at: https://www.forbes.com/sites/robsalkowitz/2022/09/16/midjourney-

founder-david-holz-on-the-impact-of-ai-on-art-imagination-and-the-creative-economy [Accessed: 28 Aug 2023]

Stable Diffusion (2023), Stable Diffusion Online. [online] Available at: https://stablediffusionweb.com/

Yousif, S., Bolojan, D. (2021) "Deep-performance - Incorporating deep learning for automating building performance simulation in generative systems", In: A. Globa, J. van Ameijde, A. Fingrut, N. Kim, T. T. S. Lo (eds.) (2021) PROJECTIONS - Proceedings of the 26th CAADRIA Conference, Vol. 1, Chinese University of Hong Kong and Online, Hong Kong, China, 29 March - 1 April 2021, pp. 151-160. https://www.doi.org/10.52842/conf.caadria.2021.1.151