Title

Survey of Generative AI in Architecture and Design

Author

Hakimshafaei, Milad

Publication Date

2023

Peer reviewed|Thesis/dissertation

University of California

Santa Cruz

Survey of Generative AI in Architecture and Design

A thesis submitted in partial satisfaction of the

requirements for the degree of

MASTER OF SCIENCE

in

COMPUTATIONAL MEDIA

by

Milad Hakimshafaei

March 2023

The Thesis of Milad Hakimshafaei
is approved:

____
Angus G. Forbes (Chair)

____
Adam M. Smith

____

Peter Biehl

Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

Abstract

Survey of Generative AI in Architecture and Design

by

Milad Hakimshafaei

This thesis gives insight into how architects and other designers can make use of generative AI tools for generating novel conceptual designs to assist in the creative process. To do this, I examine the potential uses of generative AI platforms such as Midjourney, DALL-E 2, and Stable Diffusion in architecture and design. I study the use of these generative AI platforms in producing complex designs that can be compared to those generated by existing architecture generative tools. The method used for demonstrating the capabilities of the mentioned AI platforms is to use the same prompts for each platform and run multiple tests to make a more accurate comparison of results. A number of tests are conducted, ranging from the design of buildings and architectural spaces by including factors such as traditional architectural styles, complex forms from nature, and the combination of famous architects' styles. Therefore, It helps to test how well AI can handle complex ideas that are difficult for humans to envision and difficult to implement using algorithmic tools such as Grasshopper. As part of the thesis, I survey machine learning architectures used in image-based generative AI and provide comprehensive examples of how the most

popular AI tools (Midjourney, DALL-E 2, and Stable Diffusion) translate speculative concepts into novel images.

To my family and friends

Acknowledgments

First, I would like to thank my advisor, Prof. Angus Forbes, for believing in me, for mentorship and support, and Prof. Adam Smith for his valuable insights and suggestions. I would also like to thank Dr. Oskar Elek for supporting and helping me in various stages during my academic journey. In addition, I'd like to extend my appreciation to all of my graduate school cohorts and faculty who supported me throughout this work.

1 Problem Statement

The conceptual design study is an important stage in the architectural process. In architectural design, various phases are involved, from generating initial design ideas to developing the design and preparing construction documents [90] [91]. According to a study by Marjanovic. D, et al, and Shamsudeen. Musa [73] [74], the majority of construction waste is the result of errors in the design phase. Increasing the quality of the earlier stages of the design process could lead to a more well-thought out design concept and reduce the amount of waste produced. During the conceptual design phase, the design tools (CAD) are very useful when designers are iterating over initial ideas and modeling and visualizing them in order to better picture their ideas and present them to clients. With the advancement of technology and the development of advanced software, designers were able to expand their imaginations and develop more complex and novel designs. For example, a relatively recent trend in architecture is to take inspiration from organic forms and behaviors in nature. These advanced tools use the latest approaches to algorithmic design solutions, which are called parametric or generative software. Due to the complexity associated with creating new generative algorithms, these algorithmic tools are difficult and time-consuming to use for conceptual design, although they are extremely powerful for modeling and visualizing design ideas. By using these algorithmic tools during the conceptual design phase, the designer is faced with the complexity of defining algorithms that are not necessary at this stage, as only conceptual results are required. Nevertheless, this is the most common approach used by architects from the beginning of the design process to produce complex designs.

Recent advances in artificial intelligence [52] have led to successful achievements in the area of generative design and producing images that are difficult for humans to imagine. New technologies can provide architects and designers with new opportunities to improve existing design approaches. During the past few years, GANs have been shown to be successful enough to attract designers who want to use them as part of their design process. There are a few examples, such as generating and analyzing novel architecture plans from existing datasets [71], recognizing and designing building facades [72], and ArchiGAN [92]. ArchiGAN is one such system that enables designers to generate a variety of architectural floor plans based on inputs from users. Although this is a direct and helpful application of AI in design, few designers chose to explore it. For instance, In the case of ArchiGAN focuses on floor plans rather than on broader concepts. Also, very few designers possess sufficient knowledge of machine learning to study, and GANs require this knowledge. The emergence of new generative AI platforms such as Midjourney [67], Stable Diffusion [52], and DALL-E 2 [62] has received positive feedback from architects. Since these platforms require text and descriptions to produce designs, everybody with a basic understanding of computers began to use them. By observing recent architectural events and conference topics, It's evident that they highlight the importance of addressing AI in architecture. For instance, based on the “Digital Future” talks [76], which gather most of the renowned architects (computational designers) who are actively involved in the creation of advanced and novel designs, these AI platforms have been their main focus.

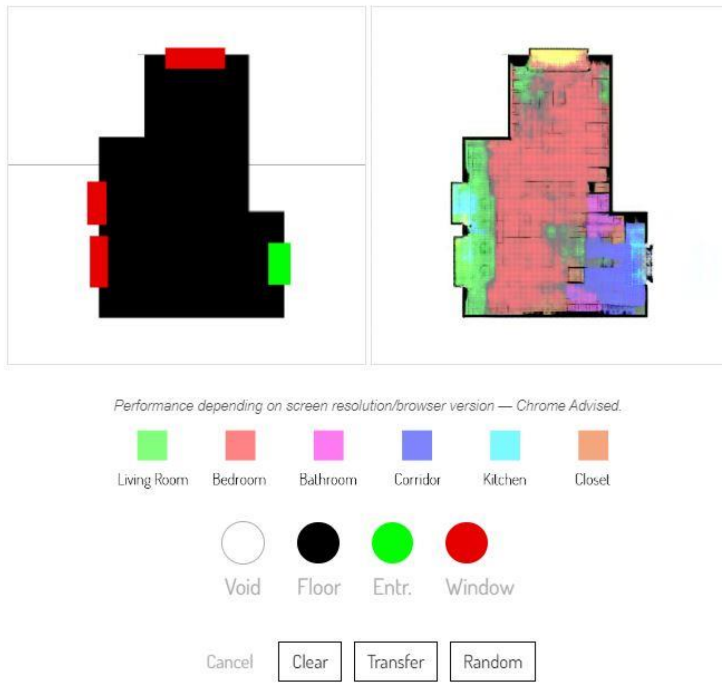


Fig 1: An experiment of generating a floor plan by ArchiGAN [92]

In my experience as an architectural designer[1], these AI tools have great potential not only because of their ease of use but also because they produce results that can sometimes be better than those generated by parametric/algorithmic architecture software. In the early stages of design, architects deal mainly with vague visual concepts or architectural-related keywords such as "cubic residential building" or "glass and metal materials" that new AI platforms are suitable to use. It is, of course,

[1] I have an educational and professional background in architectural engineering and digital architecture (computational design). As a designer, I have experience with residential and cultural buildings as well as research and development of complex form-finding studies using advanced design tools.

possible to create different design options using parametric design tools, but this is a time-consuming and difficult process. In addition, the AI is not only creating architectural forms but also adding lighting and materials, which would otherwise require the use of multiple modeling and rendering programs.

Throughout my architectural career, I have used algorithmic tools to create a variety of projects with a focus on form complexity and variety of designs, for example, involving corals in the design of residential buildings, placing brick designs on curved surfaces, and other projects which require the use of advanced design tools to simulate and imagine. Having experimented with existing architectural generative tools for many years, and now that generative AI platforms have emerged, I believe that the majority of these complex projects could be carried out more efficiently. By producing a variety of initial design proposals in a short period of time, I can then determine which specific proposal has the potential to be developed further.

Considering the variety of AI platforms available to designers, one of the areas that are lacking is an investigation of each platform focusing on a specific design problem. This enables designers to gain a deeper understanding of each generator's capabilities. **Thus, the study seeks to determine which platform is capable of understanding input from designers and incorporating all input into output with the least amount of effort and prompt engineering**. Two different groups are targeted in this thesis. Those in the first group are computational designers working in the architecture industry and testing generative AI platforms. There is a benefit to this group in studying the results generated by each AI platform in order to determine which one can be more efficient in the design process. In the second group, there are researchers who are interdisciplinary and are interested in exploring the possible application of each technology in a variety of fields. It is the purpose of this thesis to demonstrate the potential of AI in design and to present the strengths and weaknesses of each AI platform in terms of design.

To sum up, I survey both experiments by architects and some of the main tools and technologies that are now widely available for image generation and explore how these tools can be applied to architectural design and, in particular, concept generation. As an experienced architect, I provide a thorough exploration of the concept generation process using Midjourney, DALL-E 2, and Stable Diffusion, showing the strength and weaknesses of these tools for certain types of concept generation. The focus is on complex design ideas that can be done by architectural generative/parametric tools to highlight AI's potential better and provide inspiration to architects interested in incorporating generative AI tools into the concept design process.

2 Architecture and Digital Design

The advancement of computer technologies over the late 1980 and early 1990s encouraged the architecture industry to use digital tools to design complex forms, which became common and spread widely among architects' designs. These tools were not only used for the design process but also for the production of drawings that helped draw more efficient forms and assist in producing information for the construction process. Complex forms were always interesting for architects to design, which were done differently in the past without using digital tools. For instance, Antoni Gaudi, the Catalan architect, used to design computationally due to the complexity of his form derived from nature but with a non-digital approach. He made a physical model to investigate each form's structural functionality in the entire building over time, which became his tool to test his design in action. However, most of these models and drawings have been destroyed or vanished over time. His experiments are a significant inspiration for developing digital tools to simulate the behavior of real-world scale building and its relevant experiments [1].

After the invention of digital design tools in the late 80s, many architects have significantly influenced the development of these tools because of their unique design style and the need for more advanced tools and approaches. American post-modernist architect Peter Eisenman [2] played a vital role in using these digital tools in his designs as he manipulated blocks and grids generated through abstract operations. His projects are considered some of the first to use computers to compute design outputs. An

American architect Frank Gehry is one of the leading architects who influenced the use of digital tools during the 1990s [3]. His approach to design was an iterative process of physical model building and 3D modeling over years of investigation. His design process starts with a physical model, producing 3D modeling out of it, then creating a 3D physical model and manipulating it, and capturing the model with the 3D scanner to edit it with the digital model for the next steps. It used to be considered a mix of analog and digital techniques that were deployed over many years. In order to facilitate the process of digital design and fabrication process, Gehry's team developed a tool called CATIA [4], which helped simplify the design process and send the information to manufacturing machines. This tool was later developed with a different name called "Digital Project" and was released publicly for other designers to be able to quickly design and fabricate as complex forms as Gehry's team could generate [5].

The complexity and novelty of forms architects aimed at and the development of digital tools in the late 2000s led to a new era of design in the form of parametric architecture by inventing algorithmic visual scripting environments [6]. These are visual and node-based programming environments to create generative algorithms to produce 3D complex geometry. Although these tools are beneficial for structural analysis, form finding, simulation, and optimization of forms, they have a learning curve for designers since advanced computer programming knowledge is required. On the other hand, the complexity of forms and buildings designed by parametric approaches are so costly in production that few, except rich and famous architects and teams, such as Zaha Hadid, can afford them [7]. One of the main reasons is that parametric design is at its early stage as it’s less than 20 years old; consequently, manufacturing tools, architects' knowledge, and design procedures are not fully adapted to this advanced design approach.

2.1 Generative architecture tools

There is various software providing algorithmic approaches that architects use for different purposes. Some of the most popular ones include Houdini, Dynamo, and Grasshopper. As Grasshopper is extremely popular among architects, this chapter mainly focuses on it.

2.1.1 Dynamo

According to the study by Wang. Jie, et al., [8] [9], Dynamo is a visual programming platform that is built on top of Revit. Automating tasks, creating parametric models, and creating custom functionality with Revit is possible. A major benefit of Dynamo is its ability to automate repetitive and tedious tasks. In the case of architecture, Dynamo can be used to automate the creation of components such as walls and floors with specific parameters, such as their size, materials, etc. By doing so, the design process can be greatly accelerated, and errors can be reduced.

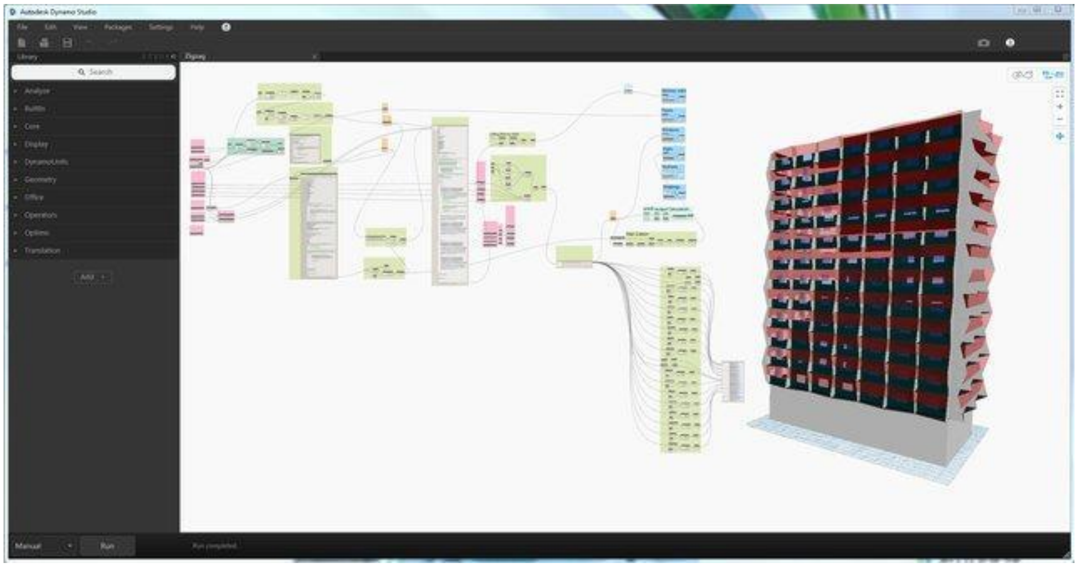


Fig 2: Dynamo graph and the geometry in Dynamo interface [10]

Dynamo also has the capability of generating parametric models. Dynamo can be used by architects and designers to create custom algorithms that generate different building designs based on specific inputs, such as the number of floors. Due to Dynamo's parametric nature, users can also easily make changes to the design and see the changes in real time, facilitating the design process.

Dynamo is also capable of performing a wide range of computational analyses. Building performance can be analyzed using Dynamo, such as energy consumption, natural light levels, and structural stability. Aside from generating different design options and optimizing the design for specific performance requirements, Dynamo can also generate custom algorithms. A number of other software platforms can be integrated with Dynamo, including Revit, Rhino, and Excel. Through Dynamo, architects and designers can extract data from one software and use that data to drive design in another. The data can then be transferred faster and more efficiently between each software, allowing architects to take advantage of the best features of each program.

2.1.2 Houdini

According to the "Visual Dataflow Modelling - Some Thoughts on Complexity " [11], which was published in 2014, SideFX Houdini software has wide applications in the architecture industry. This software is an ideal tool for architects and architectural visualizers because it is capable of creating realistic simulations, has a parametric workflow, and has a node-based procedural system.

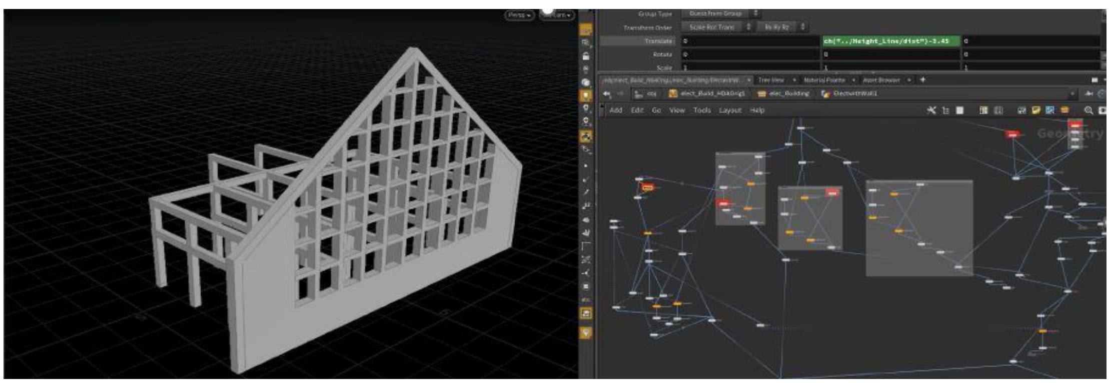


Fig 3: Reconstruction of a building from the WWII concentration camp in Croatia using procedural modeling [12]

One of the ways in which architects use Houdini is to simulate real-world conditions in a virtual environment. Houdini can be used by architects to simulate natural elements such as water, wind, and fire using physics-based simulations. An architect can use simulations to test a design's performance under extreme conditions, like a heavy rainstorm or high winds, in order to identify potential problems and customize the design before construction begins.

Architects also use Houdini to create realistic architectural renderings. This software allows users to create photorealistic renderings of buildings and environments, which can be used to showcase designs to clients and stakeholders. It is also possible to use Houdini during the design process itself. The parametric workflow of the program allows architects to experiment with a variety of variations of a design quickly, as well as make non-destructive changes to the design. Consequently, the final design can benefit from an iterative design process.

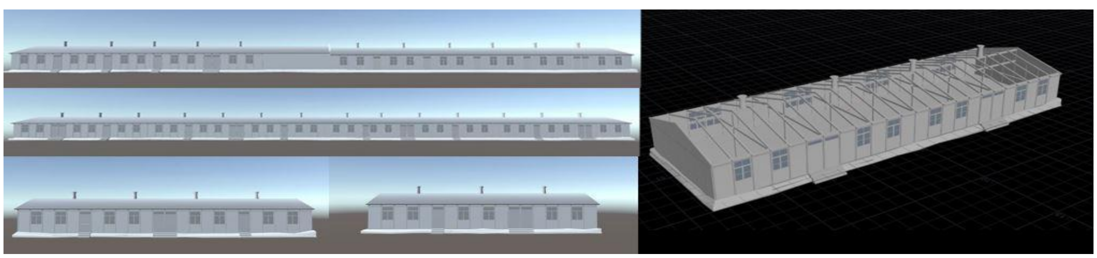


Fig 4: Procedural generation of a building from the WWII concentration camp in Germany [12]

Houdini is also equipped with a set of motion graphics tools, which can be used to create animated architectural walkthroughs, fly-throughs, and presentations. Such animations are an effective means of communicating design concepts to clients and stakeholders, allowing them to gain an understanding of the space on an entirely different level than that which can be conveyed through static renderings.

It is worth noting that Houdini is an extremely versatile tool that can be used by architects to produce photorealistic renderings, simulate real-world conditions, and perform design exploration and motion graphics on a single platform. Due to its node-based procedural system, it can provide architects with high degrees of flexibility and control, allowing them to streamline the design process and improve the final product.

2.1.3 Grasshopper

Grasshopper is widely used in architecture and design to create complex geometry, perform computational analysis, and automate tedious processes. This tool allows you to quickly generate and manipulate parametric geometry, which is a key advantage of using Grasshopper. Architects and designers can explore a wide range of design options and quickly iterate on ideas this way. In addition, Grasshopper allows users to easily modify designs and see the results in real-time, which greatly speeds up the creation process [13].

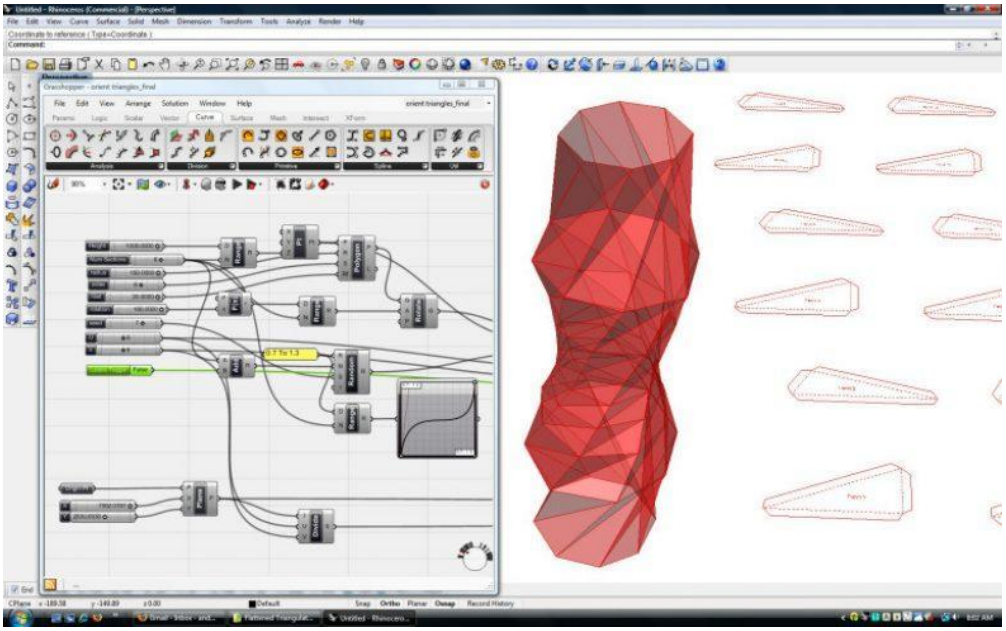


Fig 5: Grasshopper interface [14]

Computational analyses are another powerful feature of Grasshopper. Grasshopper can also be used by architects to analyze the energy consumption and natural lighting levels of a building, as well as the structural stability. Furthermore, Grasshopper can also generate different design options or optimize a design for particular performance criteria using custom algorithms [15].

As an example, Grasshopper is used for advanced computational design investigation purposes by many architects and institutes, such as Achim Menges, head of the Institute for Computational Design and Construction at Stuttgart University, whose focus is on the generation of a form and its materialization "morphogenetic design" that inspires natural forms and leverages design and fabrication [16]. Algorithmic tools have the potential to deal with and generate complex design proposals that have been used in different research projects, such as computational design solutions for high-rise building facades [17], modeling and simulation of complex, hard-to-imagine and visualizing materials in the case of exploration of weft-knitted spacer fabrics [18], and focusing on the material and structural properties of a building to better optimize results derived from the early stages of the generative design and its impact on the entire project by developing a tool for algorithmic visual scripting tools [19].

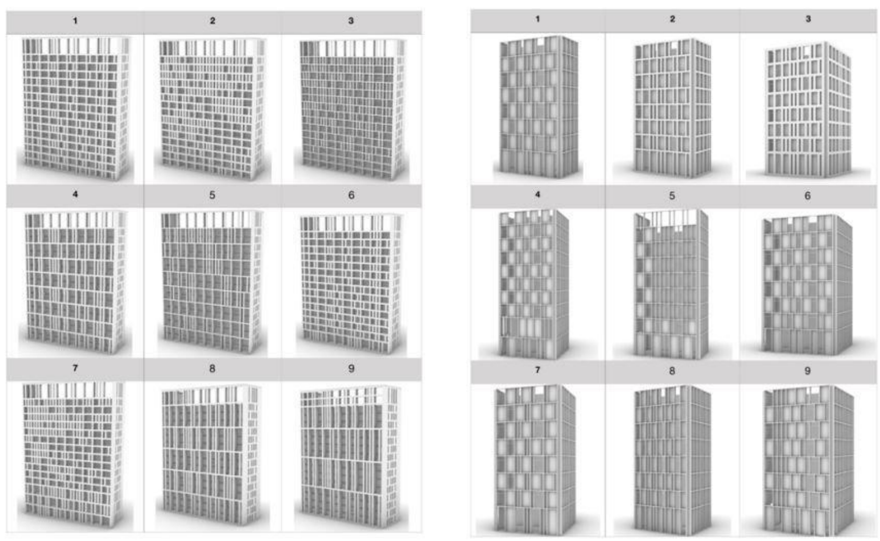


Fig 6: Generated facade designs in different styles using Grasshopper [17]

The rapid exploration of design options, detailed analysis, and automation of tedious tasks made Grasshopper an indispensable tool for many architects and designers. With its ability to adapt and make necessary changes to the design at any stage of a project, it also allows architects and designers to generate detailed construction documentation quickly. It is important to note that Grasshopper is not a standalone software but an add-on for Rhino 3D, a 3D modeling program.

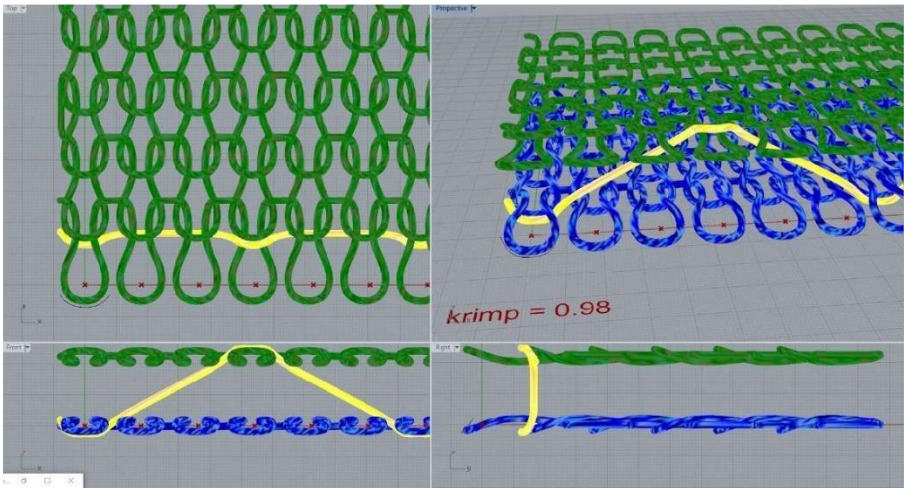


Fig 7: Fabric design by Grasshopper [18]

2.2 Discussion

Throughout this chapter, I have discussed a few common and powerful architecture generative tools. There are common features among these tools. To begin with, they are typically used for complex projects that are difficult to imagine and model using non-algorithmic software. Furthermore, all of them require advanced knowledge of algorithmic approaches, which is not a skill that all designers possess. Consequently, these tools are used by computational designers, which are a small group of professionals within the architecture industry. To better understand generative AI's potential and determine if they are superior to existing architecture tools, I discussed these tools (Dynamo, Houdini, and Grasshopper) and their applications in architecture.

3 Generative models

A generative model is an algorithm that generates new samples of data similar to a training dataset from previously unseen samples. In deep learning, generative models are usually implemented using neural networks trained to generate data samples that resemble the training data. An idea behind generative models is to capture the inner probability distribution that generates a class of data so that similar data can be generated. A variety of tasks can be performed with this, such as fast indexing and retrieval of data. Many fields and problems have been addressed by generative models, including visual recognition tasks, speech recognition and generation, natural language processing, and robotics [20] [21].

3.1 Types of Generative Models

Generative models come in a variety of types, including Autoencoders, Generative Adversarial Networks (GAN), Deep Boltzmann machines (DBMs), and Diffusion models [21].

3.1.1 Autoencoders (AE)

According to the study by Bank. Dor, et al, and Baldi. Pierre [22] [23], an autoencoder is a form of the neural network utilized for feature learning and dimensionality reduction. They are made up of an encoder and a decoder. In order for the decoder to recreate the original input data, the encoder converts the input data to a lower-dimensional representation or encoding.

Autoencoders can be considered a particular class of data compression algorithm. The encoder reduces the dimensions of the input data, while the decoder uses this reduced-dimensional representation to reassemble the original data. Although autoencoders are designed to reduce the dimensionality of the encoding while still preserving as much information as feasible in the encoding, they differ from conventional data compression methods in that they are not intended to minimize the loss of information.

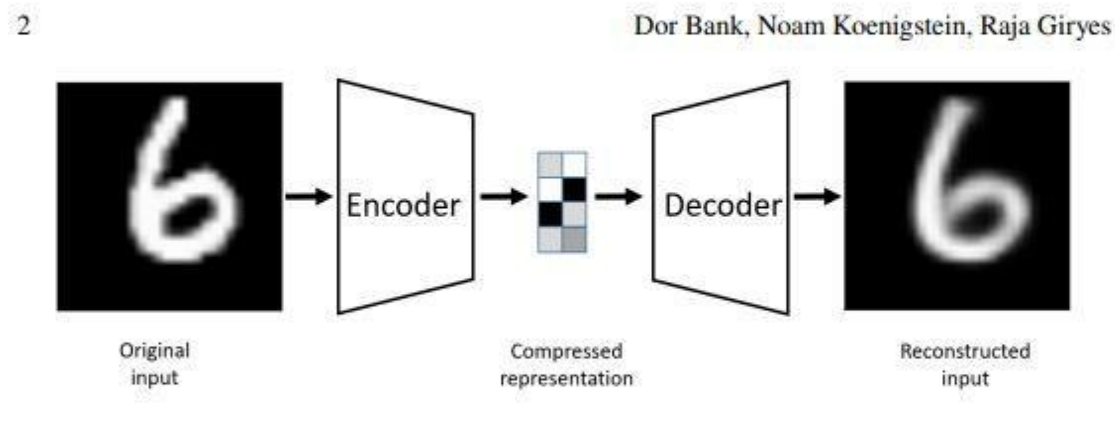


Fig 8: An autoencoder example. The input is encoded to a compressed representation and then decoded [22].

Both supervised and unsupervised methods can be used to train autoencoders. An autoencoder is trained to rebuild a labeled dataset, such as an image dataset with labels for the many classes of objects shown in the photos, in supervised training. In unsupervised training, the autoencoder is taught to rebuild a dataset that has not been labeled, such as a set of unlabeled photographs.

Convolutional autoencoders, denoising autoencoders, and variational autoencoders are only a few of the different types of autoencoders. These variations are applicable to a number of tasks, such as image denoising, picture production, and anomaly detection. In this section, we only focus on Variational Autoencoders as the most popular model.

Variational Autoencoders (VAE) are a sort of generative model that can be applied to a number of tasks, such as the creation of text, images, and representational learning. At its most basic level, VAEs function by condensing the data from an input data point into a small latent space, then using that latent representation to recreate the original data point. This procedure is comparable to that of a conventional autoencoder, but VAEs are more potent and versatile since they impose an additional restriction on the latent space.

Because VAEs are probabilistic models, they can be used to estimate the underlying distribution of data. This is achieved by defining a prior distribution over the latent space, then using the data to learn a posterior distribution that approximates the true underlying distribution. By sampling from the latent space, the VAE generates new data points and reconstructs the original data using the decoder.

VAEs have the advantage of learning compact and meaningful representations of data. As VAEs are trained to reconstruct the original data points as accurately as possible, they are motivated to learn a latent space that captures the most significant features. There are several contexts where this can be useful, including data visualization, dimensionality reduction, and feature learning [24] [25] [26].

3.1.2 Generative Adversarial Networks (GANs)

A generative adversarial network (GAN) is a machine learning model that generates new, synthetic data samples that are similar to the training dataset. In these models, two components are involved: a generator and a discriminator. The generator produces synthetic samples, whereas the discriminator distinguishes between synthetic and real samples. There is a two-part training process for the generator and the discriminator. A synthetic sample is first generated by the generator, and then a real sample from the training dataset is fed to the discriminator along with it. Discriminators try to distinguish synthetic samples from real samples and give feedback to generators.

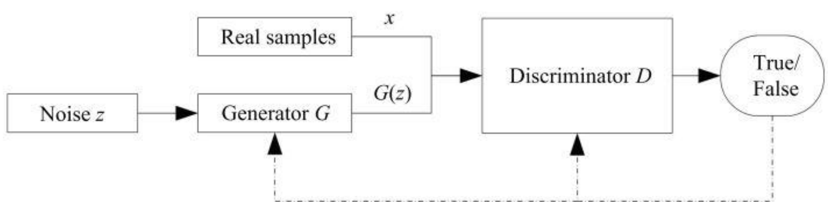


Fig 9: Structure of GANs [28]

Through this feedback, the generator improves its ability to produce synthetic samples that are similar to real ones. By repeating this process, the generator will be able to produce synthetic samples that cannot be differentiated from real samples by the discriminator. An array of applications have been achieved with GANs, including generating synthetic images, text, and audio. Additionally, they have been used to generate realistic 3D models and animations, as well as to improve image and video translation tasks.

A major advantage of GANs is their ability to generate synthetic samples that closely resemble real ones, even when the training dataset is relatively small. Their limited availability or difficulty in collecting more data makes them a useful tool for tasks where data is scarce or expensive to collect [27] [28].

Generic adversarial networks (GANs) come in several different types, each with unique characteristics. The following are some of the most common types of GANs [27]:

1: The simplest GANs are vanilla GANs, which consist of a generator and a discriminator network. Generators produce synthetic samples, while discriminators distinguish between real and synthetic samples [29].

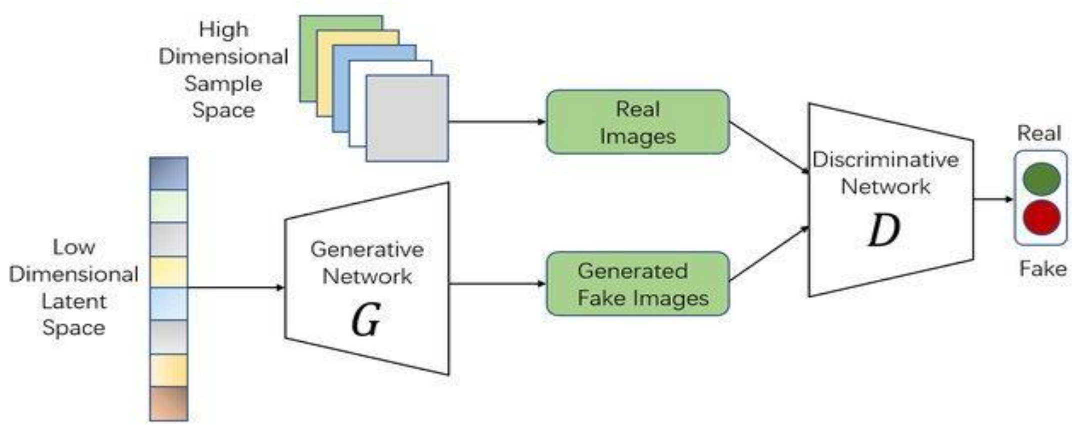


Fig 10: The architecture of vanilla GANs [29]

2: DCGANs (Deep Convolutional GANs): These GANs use deep convolutional neural networks as their generators and discriminators. A common application of these technologies is the generation of images, including faces, landscapes, and animals, which have been successfully generated with high quality [30].

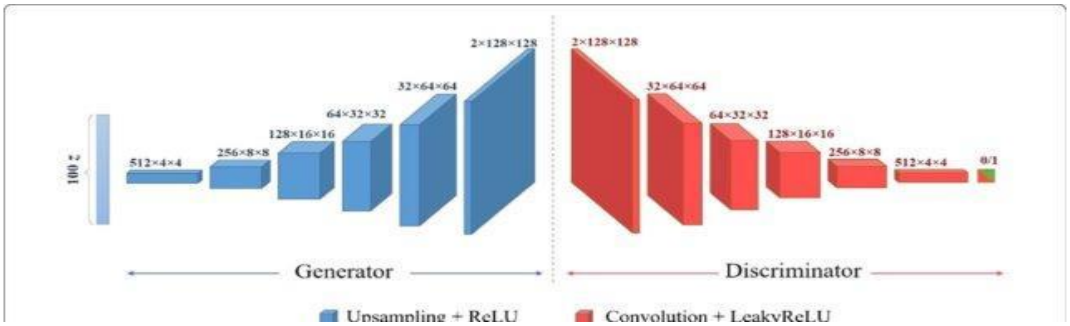


Fig 11: Deep convolutional generative adversarial network (DCGAN) architecture [30]

3: The Wasserstein GAN (WGAN) is a type of GAN that tries to generate higher-quality outputs by measuring the distance between the generator's output and real data. High-resolution images and videos are often generated using WGANs [31].

4: InfoGANs: These GANs learn the distribution of data by using an additional network called the Q network. By adjusting the input to the Q network, InfoGANs can generate synthetic samples that are more controllable and can be manipulated in specific ways [32].

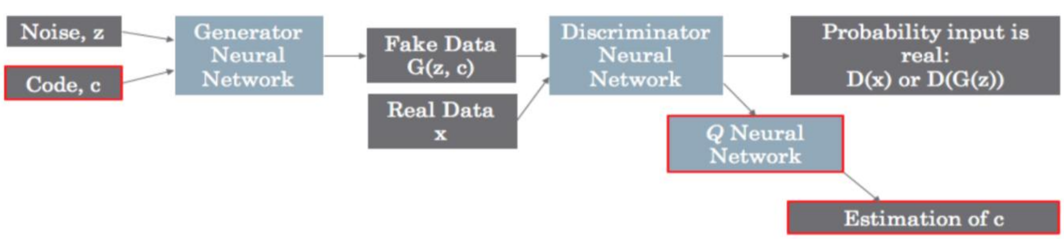


Fig 12: InfoGAN architecture [33]

5: CycleGANs: These GANs are used for image-to-image translations and consist of two generator networks and two discriminators. The first generator network generates synthetic images from a source domain, while the second generates synthetic images based on the original domain [34].

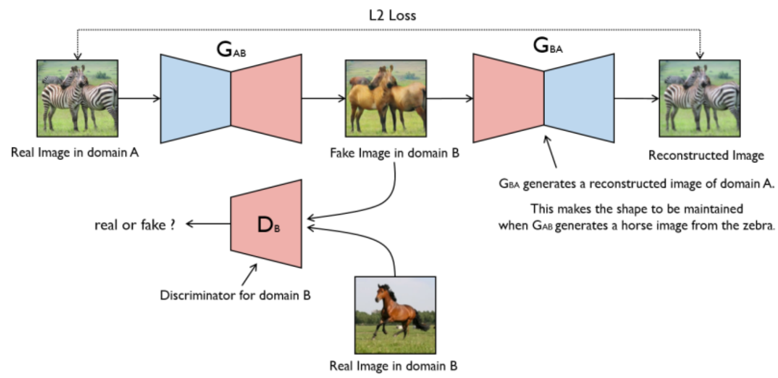


Fig 13: CycleGAN architecture [35]

3.1.3 Deep Boltzmann machines (DBMs)

Deep Boltzmann Machines (DBMs) are generative stochastic neural networks trained using contrastive divergence algorithms. The network is made up of multiple layers, with hidden units, visible units, and undirected connections between them. There are several functions that can be performed with DBMs, including dimensionality reduction, feature learning, and the generation of synthetic data. Images or natural language text distributions, for example, are particularly useful for tasks that require modeling high-dimensional data distributions.

DBMs are capable of learning complex, hierarchical representations of data, which is one of their main advantages. Data in a DBM passes through layers of hidden units that catch increasingly abstract features as it passes through the network. As a result, DBMs can generate samples that are highly realistic and diverse by capturing the underlying structure of the data. Models trained using DBMs are also highly flexible and can handle a variety of data types and tasks. The model can be trained unsupervised by only providing input data and not labeled output data, or it can be trained supervised by providing input and output data to the model [36] [37].

3.1.4 Diffusion models

Based on the study by Luo. Calvin and Yang. Ling of the comprehensive explanation of diffusion models [38] [39], these models are generative models, which generate data similar to those used to train them. The basis of Diffusion Models is that the data is destroyed by successively adding Gaussian noise, and then recovered by reversing the noise process. We can generate data using the Diffusion Model by simply passing random noise through the learned denoising process after training.

There is a fundamental difference between diffusion models and previous generative methods. Essentially, they decompose the image generation process (sampling) into many smaller "denoising" steps. By doing this, the model is thought to be able to correct itself over time and produce a decent sample gradually.

3.1.4.1 Diffusion process

Diffusion models are based on a simple concept. [40] [41] explains that the input image x0 is gradually filled with Gaussian noise through a series of T steps. This is called the forward process. In contrast to neural networks, this has nothing to do with the forward pass. Once the noising process is reversed, a neural network is trained to recover the original data. We can generate new data by modeling the reverse process. The reverse diffusion process or sampling process of a generative model is often called the reverse diffusion process.

When the noise level is low enough, sampling chain transitions can be set to conditional Gaussians. A simple parameterization of the forward process can be obtained by combining this fact with the Markov assumption:

$$q(\mathbf{x}_{1:T}|\mathbf{x}_0) := \prod_{t=1}^{T} q(\mathbf{x}_t|\mathbf{x}_{t-1}) := \prod_{t=1}^{T} \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I})$$

Where β1,...,βT is a variance schedule (either learned or fixed) which, if well-behaved, ensures that xT is nearly an isotropic Gaussian for sufficiently large T.

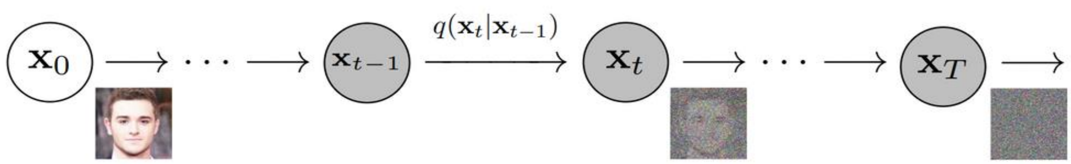


Fig 14: Forward diffusion process [42]

As mentioned previously, the "magic" of diffusion models comes in the reverse process. During training, the model learns to reverse this diffusion process in order to generate new data. Starting with the pure Gaussian noise

p(xT):=N(xT,0,I), the model learns the joint distribution pθ(x0:T) as

$$p_\theta(\mathbf{x}_{0:T}) := p(\mathbf{x}_T) \prod_{t=1}^{T} p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) := p(\mathbf{x}_T) \prod_{t=1}^{T} \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t))$$

In this case, the parameters of the Gaussian transitions are learned over time. It is important to note that the Markov formulation states that reverse diffusion transition distributions are determined by the previous timestep only.

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) := \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t))$$

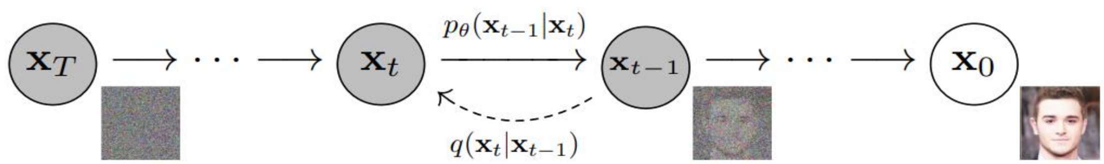


Fig 15: Reverse diffusion process [42]

3.1.4.2 Diffusion models architecture

A diffusion model architecture is deployed by U-Net. With a U-Net, inputs and outputs share a common spatial dimension and skip connections are used between encoders and decoders with corresponding feature dimensions. Generally, the input image is downsampled and then upsampled to reach its original size [41].

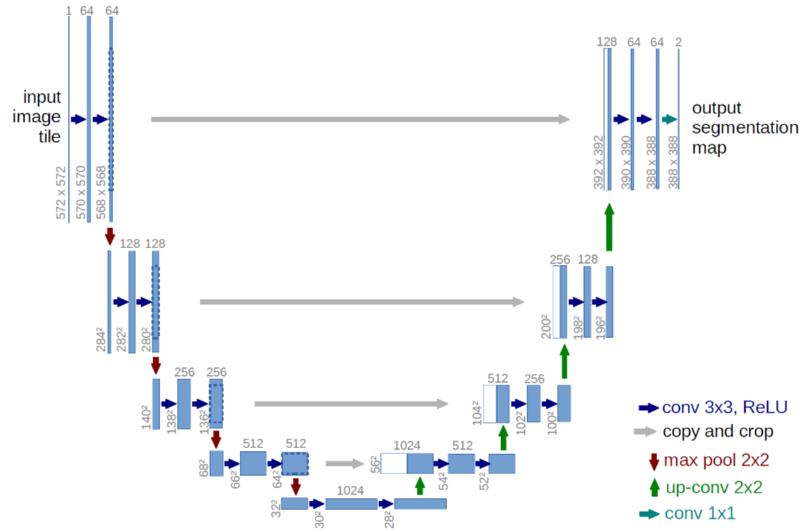


Fig 16: Architecture of U-Net [43]

3.1.4.3 Applications of diffusion models

The categorization of diffusion models is based on their applications. According to the [39] paper, these models can be used in computer vision, natural language processing, temporal data modeling, multi-modal learning, robust learning and interdisciplinary applications. This thesis focuses on computer vision and multi-modal learning, which can be used to generate images, videos, and point clouds that can be beneficial to the architecture industry.

Here are some computer vision applications of diffusion models.

3.1.4.4 Super resolution

In diffusion models, super resolution refers to a technique used to enhance image resolution. Through mathematical algorithms, low-resolution images are converted into high-resolution images. As a result, the image becomes more detailed and contains more information than the original low-resolution image. This technique is particularly useful in fields that often have limited resolution, such as microscopy imaging, satellite imagery, and medical imaging [39] [44].

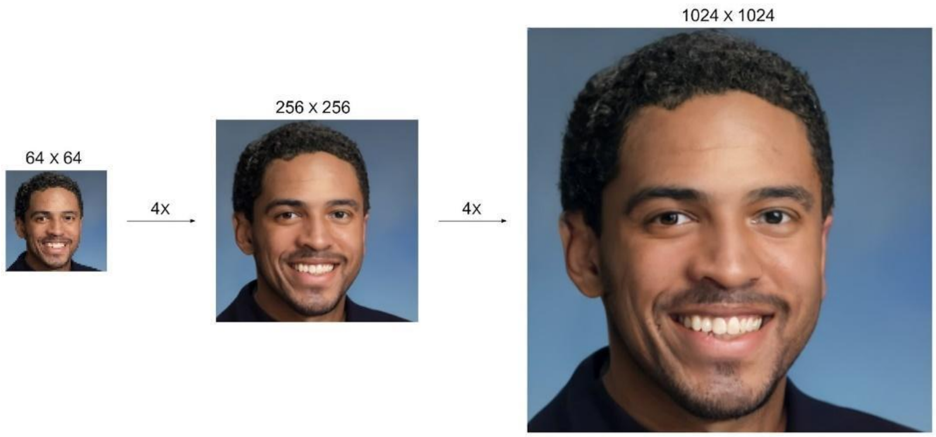


Fig 17: Upscaling in several steps [45]

3.1.4.5 Inpainting

The process of inpainting in diffusion models is used to fill in missing or corrupted parts of an image by means of the diffusion process. Mathematical algorithms are used to propagate information from the known to the unknown or corrupted parts of an image. The result is an image that replicates the original but with missing or corrupted areas filled in. Images can be restored through the inpainting of diffusion models when they have been damaged or degraded by factors such as noise, blur, or compression. Medical imaging, image processing, computer vision, and other fields where image quality is important can benefit from the technique [39].

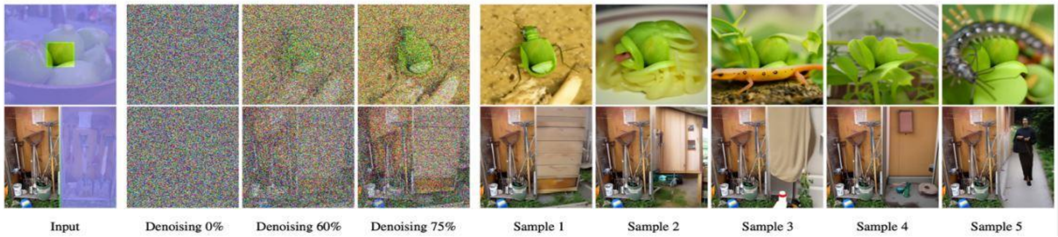


Fig 18: Image inpainting result [46]

3.1.4.6 Image Translation

An image can be translated from one domain to another using the image translation technique. Through mathematical algorithms, information is spread from the original image to the target image. As a result, a new image is created that is similar to the original, but has certain characteristics or attributes that have been modified [47].

3.1.4.7 Point Cloud Completion and Generation

Point clouds are a crucial part of the 3D representation of real-world objects. Due to partial observation or self-occlusion, scans often result in incomplete point clouds. To address this challenge, diffusion models have recently been applied to infer missing parts and reconstruct complete shapes. This technology is being used in a variety of applications, including 3D reconstruction, augmented reality, and scene understanding [39].

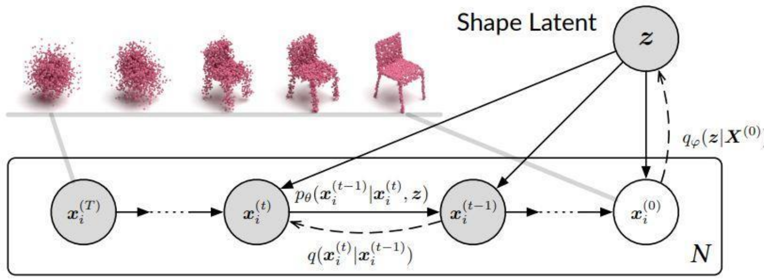


Fig 19: The directed graphical model of the diffusion process for point clouds [48]

3.1.4.8 Multi-modal learning

The term "multi-modal" is used in diffusion models to refer to the ability to generate multiple modes or patterns. Modes are patterns or characteristics that can be observed in the data. Rather than capturing and generating only one mode of data, a multi-modal diffusion model can generate and capture a wide range of variations. Application for this capability is when the data exhibited a wide range of variations and several outputs were required, such as images, speech, or text. In this section, we focus on image generation [49].

3.1.4.9 Text-To-Image generation

In text-to-image generation, machine learning is used to create images based on text descriptions. The description in the text can be input into a diffusion model, which generates an image matching the description. Creating a meaningful image from a text is a challenging task that calls for a model that can determine what the text means [50] [51].

There are a number of popular generative text to image platforms, including DALLE-2, Stable Diffusion, and Midjourney. This section explains how these systems work, as well as their capabilities.

3.1.4.9.1 Stable Diffusion

According to the study by Rombach. Robin, et al. explaining image synthesis with latent diffusion models and multiple explanations from blog posts [52] [53] [54] [55] [56] by using Stable Diffusion, you can generate images from natural language descriptions through text-to-image generation. As input, a text description is taken and used to generate an image. A diffusion model denoises random Gaussian noise step by step in order to produce an image. The disadvantage of this process is that it is slow and expensive. The introduction of stable diffusion was made to solve this problem since it relies on latent diffusion. Rather than using pixel space, latent diffusion uses a lower-dimensional latent space to reduce memory and computational costs. Latent diffusion has three main components: **Autoencoder (VAE)**, **U-NET,** and **Text-Encoder**.

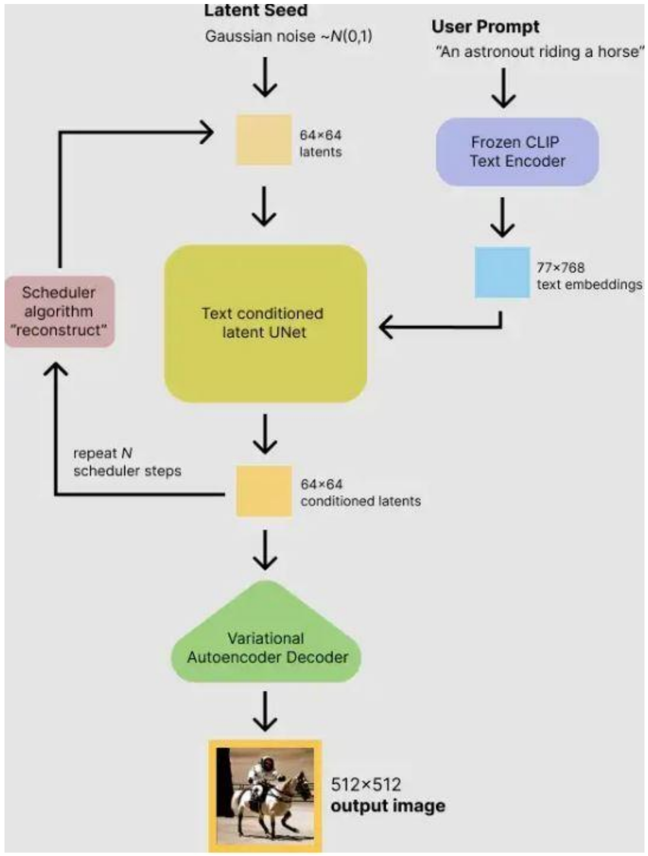


Fig 20: Stable Diffusion model [56]

Autoencoder (VAE): VAE neural networks consist of two parts: an encoder and a decoder. A latent space encoder reduces an image's dimension. From the latent space, the decoder restores the image. Latent noise is generated instead of a noisy image during training. In contrast to corrupting an image, latent noise corrupts the representation of the image in latent space. As a result, the latent space is smaller and so it is much faster.

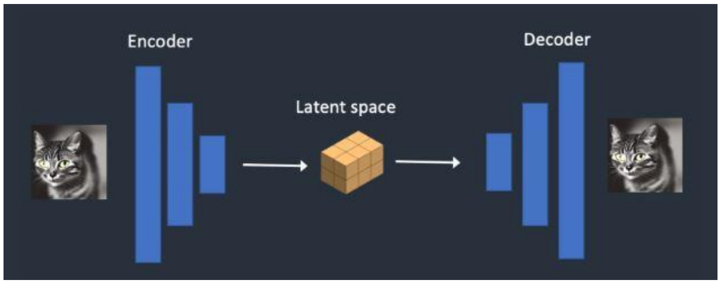


Fig 21: Variational autoencoder transforms the image to and from the latent space [53]

U-NET: As with the U-Net, the U-Net consists of both encoders and decoders, both of which use ResNet blocks. In the encoder, an image representation is compressed into a lower-resolution image, and in the decoder, the lower-resolution image is decoded back into a higher-resolution image.

It is usually necessary to add short-cut connections between the encoder's downsampling ResNets and the decoder's upsampling ResNets to prevent the encoder from losing information during downsampling.

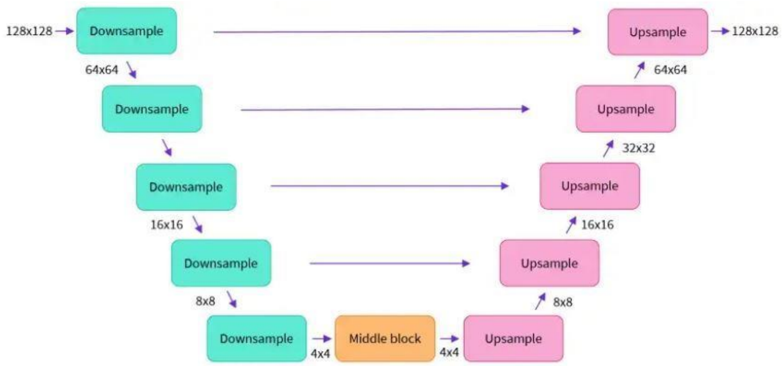


Fig 22: The U-Net Architecture [55]

Text Encoder: As part of this step, we add text to the model and define the concept of conditioning. Conditioning is the process of steering the noise predictor so we will get the image we want after subtracting from it the predicted noise. The word needs to be tokenized for the machine to understand it. Humans can read words, but computers can only read numbers. The first thing a text prompt does is convert the words into numbers [58].

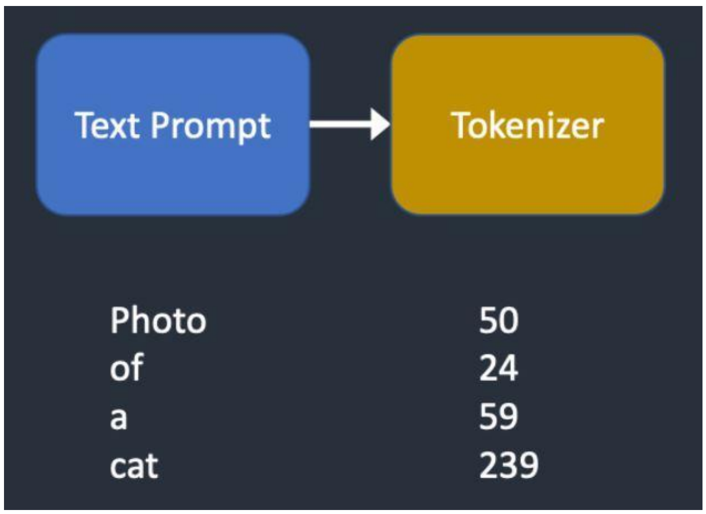


Fig 23: Tokenizer [53]

Tokenization is carried out by CLIP tokenizer. A CLIP tokenizer stands for Contrastive Language-Image Pre-training. The method uses a combination of images and text data to train natural language processing tasks like understanding languages, classifying texts, and translating them. With CLIP, the goal is to improve the representation of text data by leveraging a large amount of visual information present in images. To do this, you can train a model to predict which image will most closely resemble a given piece of text, and vice versa.

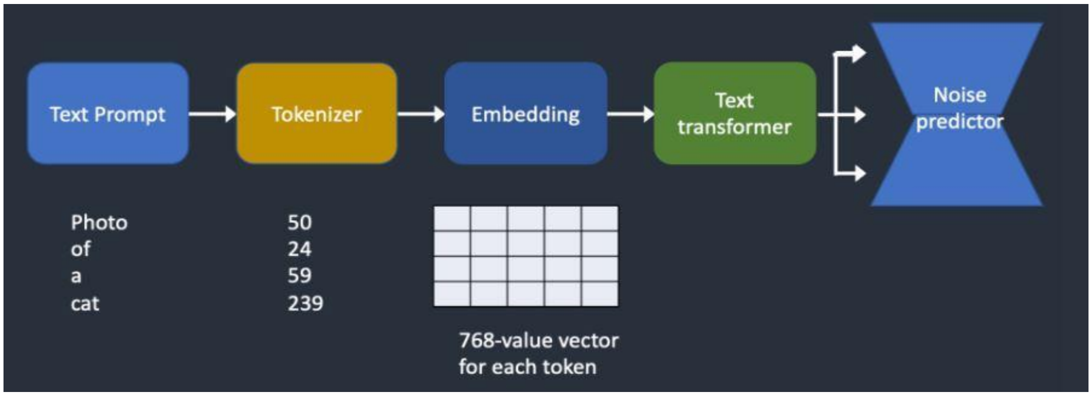


Fig 24: How text prompt is processed and fed into the noise predictor to steer image generation [53]

The next step is embedding, which describes words or tokens as continuous, low-dimensional vector spaces. Embedding aims to capture the meaning and context of a word or token so that it can be easily incorporated into other models, such as neural networks [59].

Stable diffusion platforms that are free to use: **Dream Studio**, **Mage Space**, **Playground AI**, **Dezgo**, **Neural.love**, **Stable Horde**, **Craiyon**, **Dream By Wombo**, **Night Cafe**

There are a number of tasks that stable diffusion is capable of performing, including:

Text to Image generation: Produce a variety of images based on the description provided in the text.

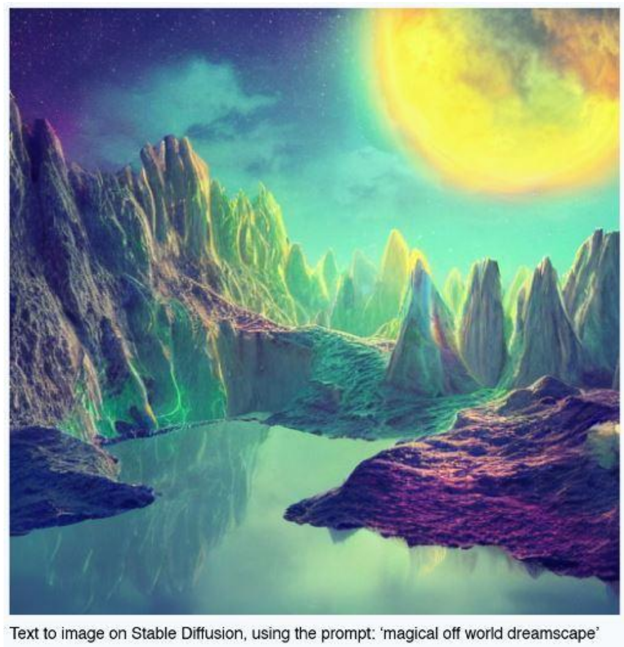


Fig 25: Text to image using Stable Diffusion [56]

Image editing: Create modified images based on user input, for example, removing objects from images or adding new objects.

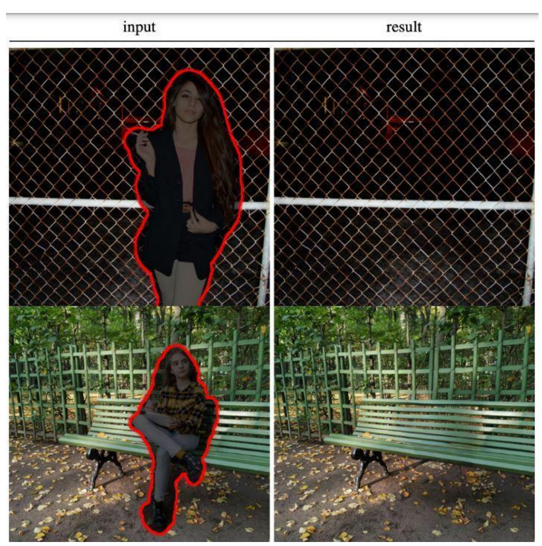


Fig 26: High resolution inpainting using Stable Diffusion [60]

Image-to-image translation: Creating new images that are visually similar to an input image, such as converting a day image to a night image or changing the colors.

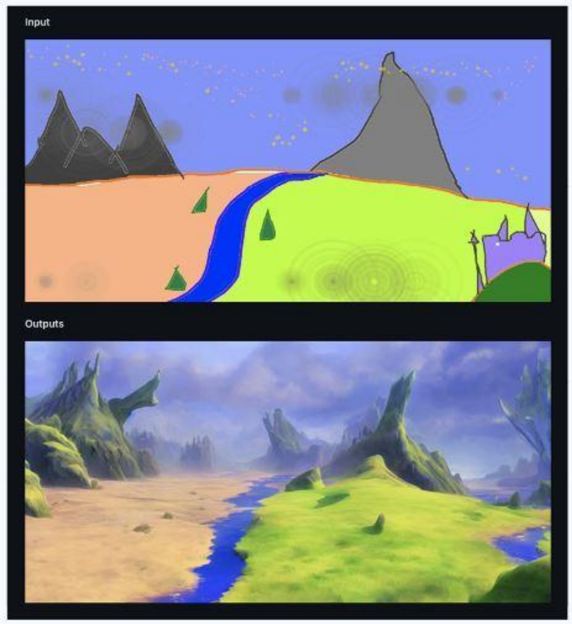


Image-to-image translation using Stable Diffusion

Fig 27: Image to image translation using Stable Diffusion [56]

3.1.4.9.2 DALL-E 2

The DALL-E 2 model replaces the Dall-E model developed by Open AI. Dall-E is a portmanteau of Wall-E (a Pixar sci-fi film) and Salvador Dali a Spanish artist known for his surrealistic style. From a text description, this model creates photorealistic images [61].

The DALL-E 2 consists of two models: CLIP and Diffusion. Given a caption and a diffusion model, CLIP generates CLIP text embeddings based on the caption. The diffusion model then generates the image embeddings based on the CLIP text embeddings. The final image is generated by decoding this embedding using a diffusion decoder [62].

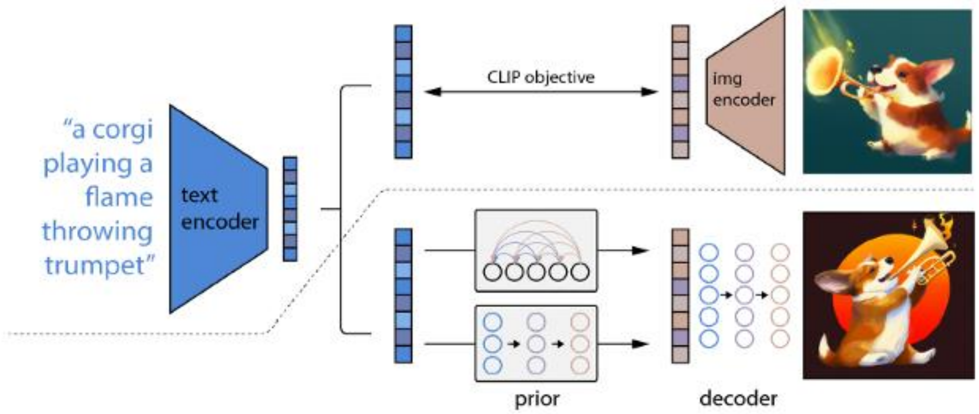


Fig 28: DALL-E 2 training process [62]

OpenAI developed a decoder for DALL-E 2 called (Guided Language to Image Diffusion for Generation and Editing). The GLIDE diffusion model is a modified version of diffusion models. By incorporating textual information, it distinguishes itself from pure diffusion models.

GLIDE: The diffusion model starts from randomly sampled Gaussian noise, so it can't generate specific images. With GLIDE, you can augment the training process with additional textual embeddings, building on the success of Diffusion Models. The image is then generated based on the text condition. GLIDE is what enables DALL-E 2 to edit images using text prompts. The GLIDE model used as the decoder in DALL·E 2 is slightly modified. In addition to text information, there are CLIP embeddings as well [63] [64] [65].

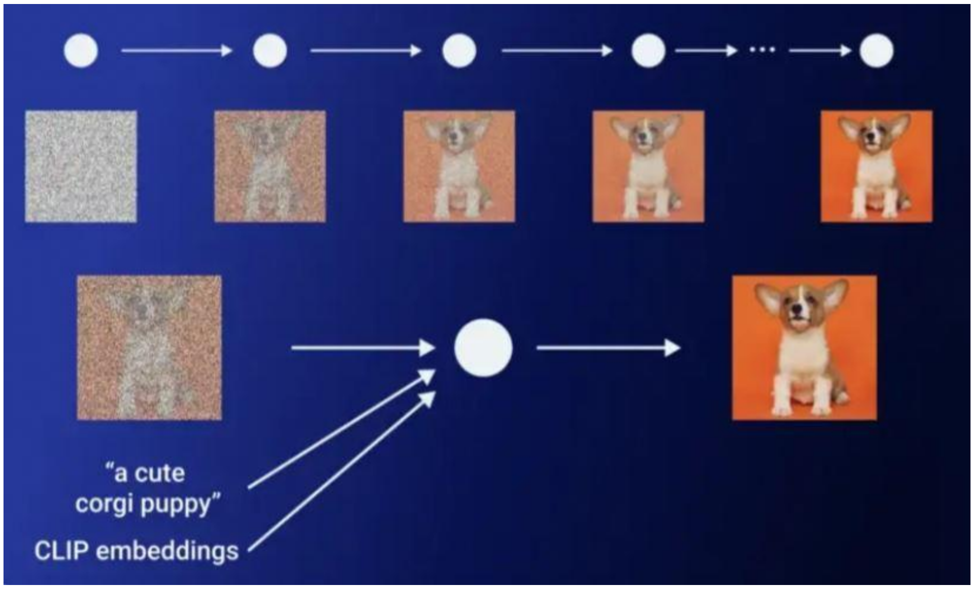


Fig 29: DALL-E 2's modified GLIDE adds the caption CLIP embeddings [64]

There are a number of tasks that DALL-E 2 is capable of performing, including:

Text to image generation: From a text description, you can create original, realistic images and art with a variety of attributes, concepts, and styles.

Outpainting: A user can extend an original image by outpainting, creating larger pictures in any aspect ratio. DALL-E 2 maintains the original image context by allowing you to enter prompts that take into account the original image's existing visual elements.

Fig 30: Outpainting by OpenAI [66]

Image editing: Edit existing images based on natural language captions. In addition to adding and removing elements, it handles shadows, reflections, and textures.

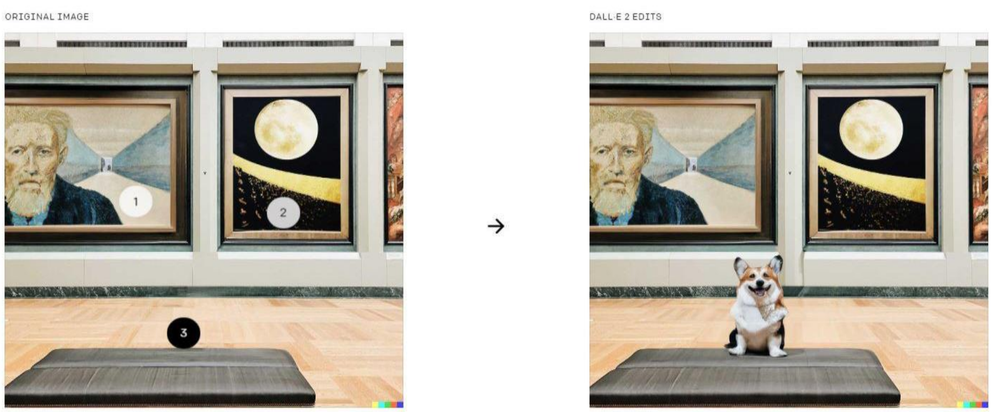


Fig 31: Image editing by OpenAI [66]

Image Variation: Create different variations of an image inspired by the original.

Fig 32: Image variation by OpenAI [66]

3.1.4.9.2 Midjourney

Similarly, Midjourney uses artificial intelligence to generate images based on prompts from the user. With MidJourney, users can create an image of any combination of things they want based on actual art styles. A special skill it excels at is creating environments with dramatic lighting that resemble rendered concept art from video games, especially fantasy and sci-fi scenes [67]. At the moment, Midjourney is only accessible via the Discord bot on the company's official Discord channel. Therefore, it is not known how the model was developed.

Text to image generation: Similar to other generative platforms, Midjourney generates images based on text prompts. Since this system runs on discord, it has more options to control the results, such as adjusting the aspect ratio, resolution, chaos, etc. [68].

Fig 33: Text to image with Midjourney, prompt: “Architects walking on a construction site and checking the project on a tablet with augmented reality “. Source: Created by the author.

Midjourney generates four images by default. It allows for generating variations from each result (V), or upscaling the targeted result (U).

Fig 34: Variation and upscaling of the generated result in Midjourney, source: Created by the author.

Image Variation: Create different variations of an image inspired by the original image.

Fig 35: Variations of a generated image in Midjourney, source: Created by the author.

Image to image: In Midjourney, it's possible to use an image with the text to generate a new result. In the following example, a house image was used as an input, and a simple text prompt, "A house building" and the generated results are close enough to the original image.

Fig 36: Generate new images (right) based on the reference image (left), source: Created by the author.

4 Generative AI and Architecture

I examined the most common architecture generative and parametric tools in previous chapters, as well as machine learning architectures used in image-based generative AI, and provided comprehensive examples demonstrating how the popular artificial intelligence tools (Midjourney, DALL-E 2, and Stable Diffusion) generate novel images based on text descriptions. This chapter aims to explore architects' opinions regarding these AI tools and their potential applications in design.

As an experienced architect who has worked for many years in this area, I hope to add my own view of AI before expressing architects' opinions on the use of AI in architecture. The conceptual design phase, which is the focus of this thesis, is where all of the ideas become visual and meaningful. Architects can present ideas using simple physical models or sketches when they are based on simple concepts. However, complex ideas require complex and powerful design tools in order to be visualized. As discussed in the previous chapters, parametric and generative architecture tools such as Grasshopper and Dynamo are frequently used to accomplish this goal. Having worked with these tools for conceptual design for many years, I can admit that since you become so focused on developing algorithms, you close your mind to new ideas. In other words, instead of considering several design options, you should focus on creating one of your ideas, which will require considerable time since the algorithm must be designed.

Architects are becoming more interested in these generative systems, but as Neil Leach says [ ], the main reason could be their simplicity in generating images since they only require natural language text as input. As mentioned, existing generative algorithmic tools that architects use need advanced programming knowledge; accordingly, simplicity in producing design proposals could be a significant advantage of AI tools. It leads to a better approach to generating various design options than algorithmic tools that require designers to design the entire algorithms for generating the model. At this stage, generative tools used by architects are more accurate in producing models. However, because this thesis focuses on the early stages of the design, it is necessary to be able to produce a variety of design proposals quickly, even if they are not technically correct or constructible. A design team prefers to see hundreds of design proposals with different concepts at the beginning of the project as opposed to different design variations from the same system that current parametric design tools are capable of producing.

I have attempted to redo the experiment in the following example (fig 35), which is part of my old studies of experimenting with different twisted towers. This design was created in Grasshopper, and defining the algorithm took a few hours. Recently, I tested it with AI (Stable Diffusion). The results are significantly superior to what I obtained with Grasshopper; in addition, it provides me with a greater variety of architectural rendering styles. If I wanted to add material and render my tower from Grasshopper, I would have had to do it with multiple 3D modeling programs, but with AI, you are able to achieve the results you desire for conceptual studies, and you can quickly evaluate various variations that are unavailable with existing architecture tools.

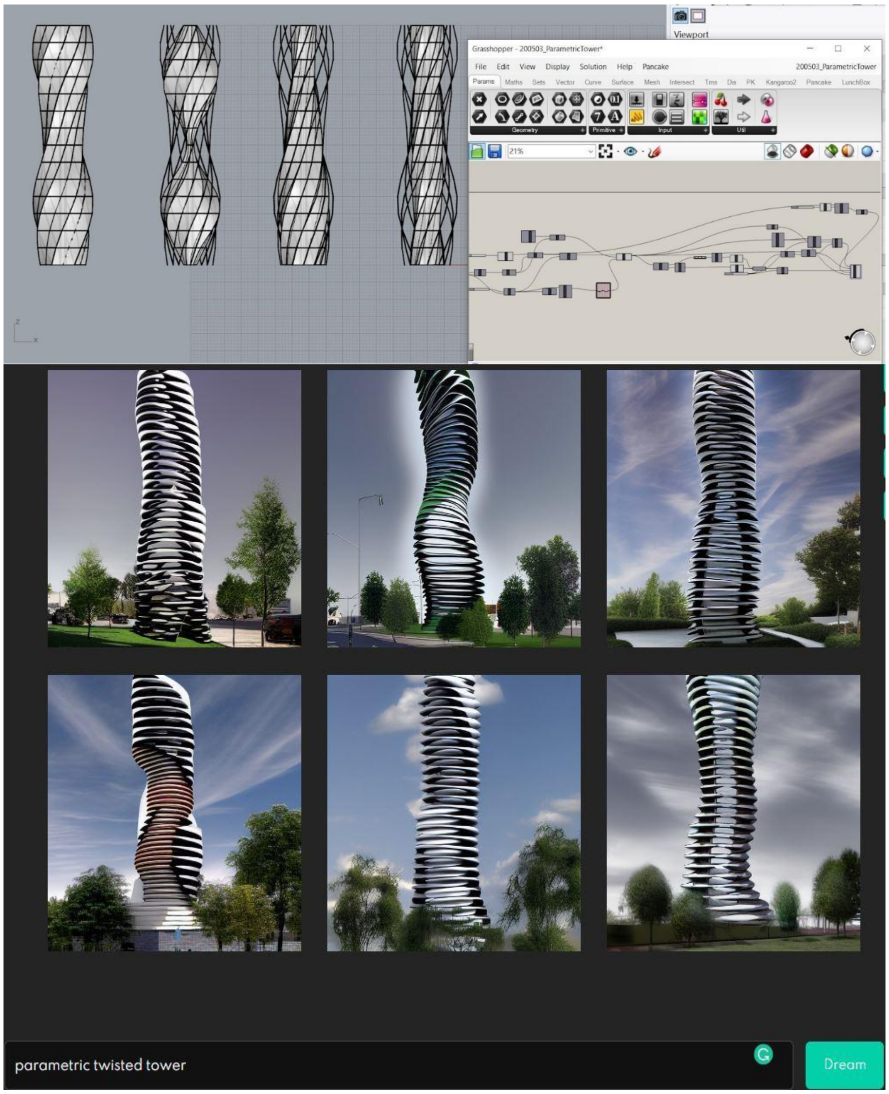


Fig 37: (top): A generative tower designed by an algorithmic visual scripting tool, Grasshopper. (bottom): Generated a variety of tower designs using Stable Diffusion (Dream Studio). Prompt: Parametric twisted tower. Source: Created by the author.

4.1 Evidence

In this section, I aim to explore how these generative AI systems can benefit architectural design. The accessibility of these AI tools and their simplicity of use encouraged architects to test them. In order to study the impact of AI on design, we rely on the experiments done by architects who have been successfully active in the last decade in implementing novel approaches in design and construction. These architects are already familiar with existing generative tools used in the design, and their feedback on using the new AI systems reflects a reliable comparison of these two generative systems.

The experiments are collected from multiple blog posts or social media platforms where these architects directly share their opinion after testing AI in design. Since this technology is new and in the experimental stage, and I could not find any published articles yet, It's difficult to find many designers sharing their experience; hence, this thesis only presents what could be found in blog posts or talks.

In this section, I briefly introduce architects I found who were already posted their feedback and opinion on using AI.

- Foster + Partners is a British architectural, engineering and integrated design practice founded in 1967, associated with the great architect Norman Foster.
- Andrew Kudless is a professor of architecture at California College of Art and founder of the design firm Matsys. He investigates material systems' emergent and integral relationships between form, growth, and behavior.

- Neil Leach is a British architect and theorist whose research shapes around critical theory and digital design.
- Behnaz Farahi is an award-winning designer and creative technologist working at the intersection of fashion, architecture and interactive design.

Foster + Partners is one of the largest architecture firms that has already explored AI tools for the conceptual design phase of projects and explained that these tools could quickly illustrate an idea or feeling that we want a particular space to evoke. They believe AI is evolving in a way that can be useful for drawing inspiration. Architect Andrew Kudless, who also tested these tools in his recent experiments, thinks they can represent a fancy version of sketchbooks that ideas can take cough form. In other words, you can dream what a project could look like. Kudless says, "architects spend hours working with computer-aided design programs to render projects representing their idea more in detail following aesthetics, but these AI tools could speed up turning your ideas into workable designs. Based on my experiments, another advantage of AI is to deal with more complex concepts that are too complicated to design. I have been interested in exploring the use of fabrics in design; however, accurately modeling and visualizing fabrics with computational design programs requires special knowledge. Accordingly, I found AI suitable and straightforward for exploring various fabric architecture designs" [75].

Fig 38: Fabric experiments with Midjourney by Andrew Kudless

Neil Leach explored different types of buildings in the style of Zaha Hadid and highlights the importance of prompt/text engineering as he finds a great future for it to encourage design in more engaging approaches to studying design. Leach says, “what we see from AI is just beginning and takes 3 to 5 years when we have a single platform doing the process of converting data to fabrication, including building regulations, cost estimates, structure, and environmental performance. I think the functionality of text-to-image systems is similar to GANs in that the system generates images, and architects act as discriminators and choose the right image generated”.

Fig 39: Style combination experiments with Midjourney by Neil Leach

Behnaz Farahi think these AI systems are beyond mood boards; instead, they are a way to imagine many possibilities. Farahi says, “the importance of imagination as the primary tool for creativity and imagination of new ideas, and generative AI tools provide the chance to imagine through text rather than visuals. Following this statement, my questions are: Does this tool challenge what it means to be creative? Does this tool shape our thought? In my opinion, since these tools can change the architectural design process by generating fast images, how will it impact the evaluation of students' design at school?" [76].

Fig 40: Fashion and building design form-finding experiments with Midjourney by Behnaz Farahi

I decided to base the results on architects' experiments and choose the ones already familar with the existing generative tools used in architecture, who also tested AI tools for design purposes. Consistently, we can rely on their feedback since they have already tested existing tools in architecture and new generative AI systems. According to the architects' experiments, they all mention AI can improve the way designers used to generate initial design ideas and include more creativity in the process. However, they still think it's too early to decide the exact impact of AI in design practices, but they think the change is undeniable not only on the professional level but also on how architecture students will be trained.

The growth of AI development in the field of generative design is quick. New features and tools are releasing such as Dream Fusion which allows producing 3D geometries derived from texts [77], Point-E for generating point clouds from text prompts [78], and a new AI plugin for Autodesk Revit software that allows producing various renders from an existing 3D model based on text input [79]. Another popular new generative system developed by OpenAI is ChatGPT, an AI language model that can communicate naturally with people. An artificial intelligence system can respond to various queries and conversations in a manner that resembles that of a human [80]. ChatGPT is being used by computational designers in architecture in order to assist them in generating complex forms and geometries using Grasshopper by having AI generate C# and Python code [81].

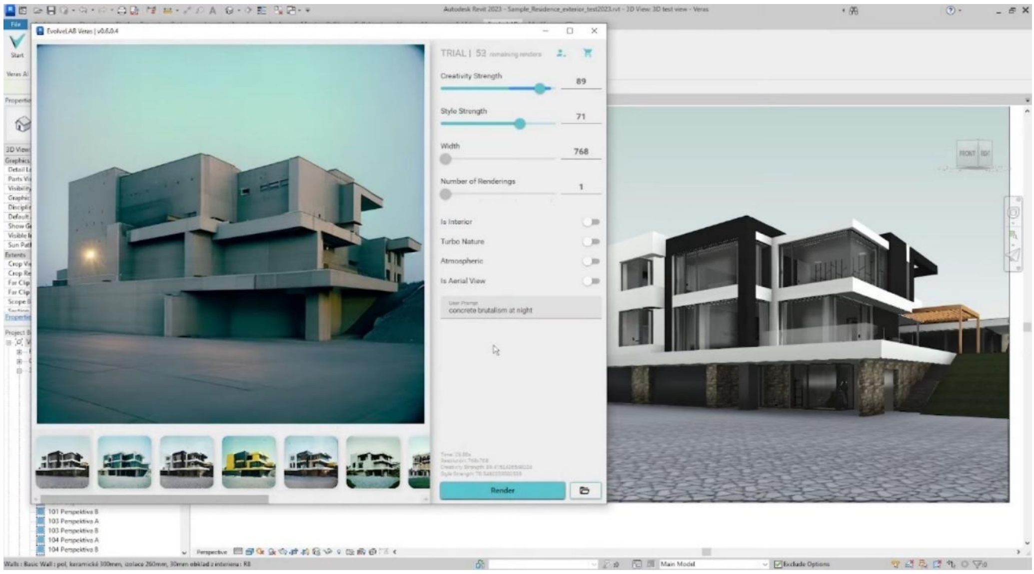


Fig 41: Produced render using VERAS AI tool inside Autodesk Revit3D software [82]

5 Generative AI Experiments

As discussed in the previous chapters, AI has various potentials for conceptual design. Following the advantages of using AI compared to existing parametric and generative software used by architects, we can point out that designers using parametric software have to define the entire algorithm themselves one at a time, and on the other hand, AI tools generate numerous examples that designers can refine, resulting in a unique approach that enables them to generate multiple ideas quickly. In addition, the software is controlled by a text interface, making it easier to use than parametric tools that use pure scripting or visual programming systems such as Grasshopper.

Although I examined a few examples of architectural experiments in the previous chapter, the experiments did not focus on the capabilities of each generator separately. I intend to explore all three new generative AI platforms, particularly emphasizing the text-to-image systems (Midjourney, DALL-E 2, and Stable Diffusion). Essentially, this study aims to study complex concepts using AI, which requires advanced parametric/generative architecture tools to design. As mentioned earlier, designers experiment with each AI tool, but what is missing is the comparison of the results generated by each AI tool when presented with the same text prompt. It allows a better understanding of the advantages of each AI tool in producing complex concepts. Nevertheless, before comparing the mentioned AI platforms and focusing on complex concepts, I first study how AI can deal with architectural terminology. I survey simple but important keywords in architectural design as part of the first experiment to determine if the results align with what an architect should expect. Afterward, I examine complex concepts using AI platforms.

5.1 Experiment 1: Conceptual Keywords to Architectural Images

The first part of this experiment focuses on testing the main terminologies that architects deal with during the conceptual design process focusing on design and visualization. My first experiment focuses on design. The **design** stage is where different forms, building types, and architectural styles are important to consider since they can shape an idea for the conceptual design. In order to define the key terms used in the design, I refer to two resources that represent comprehensive terminologies in architectural design, such as the architectural vocabulary [84] and the visual dictionary of architecture [83]. The purpose of these documents is to provide detailed terminology; however, I have chosen the most commonly used components on the basis of my experience as an architect considering the mentioned resources.

The second part of the experiment focuses on **visualization**. When producing images, it’s important to present the idea with materials and proper landscapes to better explain the idea in a more eye-catching way. Using the same method, I determined the most commonly used keywords in the visualization process based on my experience and resources [83] [85]. This part of the experiment helps test how AI understands different materials, architectural views, rendering styles, and landscapes.

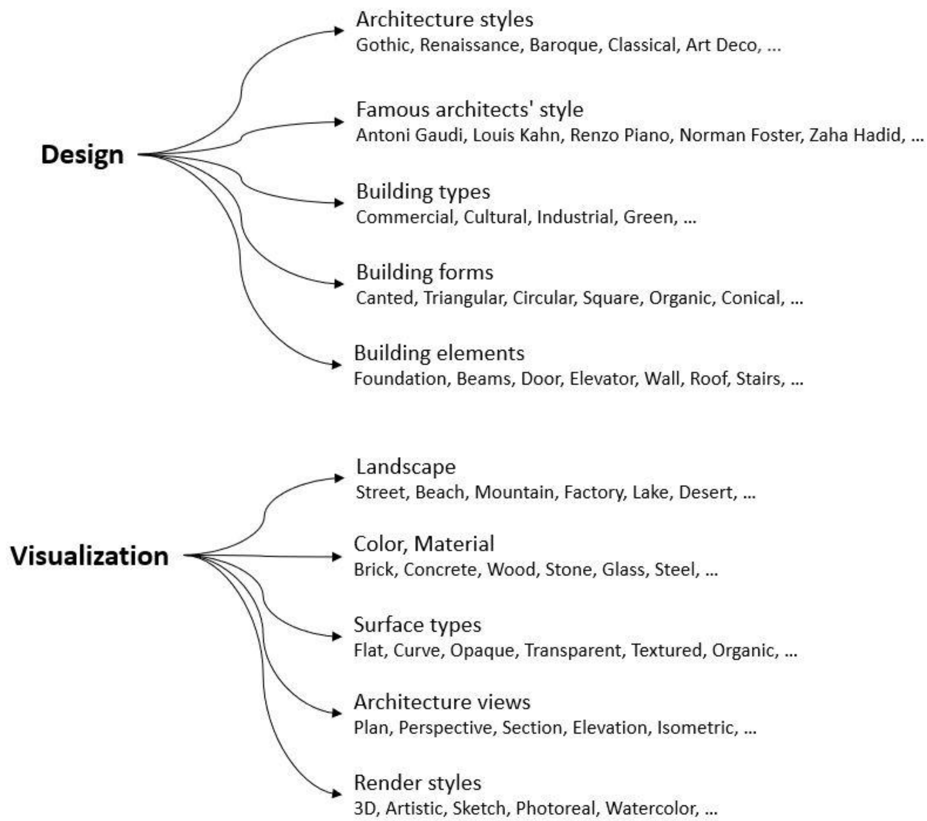


Fig 42: Design and visualization used terminologies for the AI experiments.

In order to try out these concepts and keywords, it’s required to test it with a generator that gives us more flexibility to use in terms of the time of the usage of GPU. Therefore, we chose Stable Diffusion since it’s more accessible through its google colab implementation. We didn’t change the default parameters that were set in the google colab file. The reason is to figure out what could be the first results with the default settings, which is important to study how powerful this generator is in understanding architectural keywords.

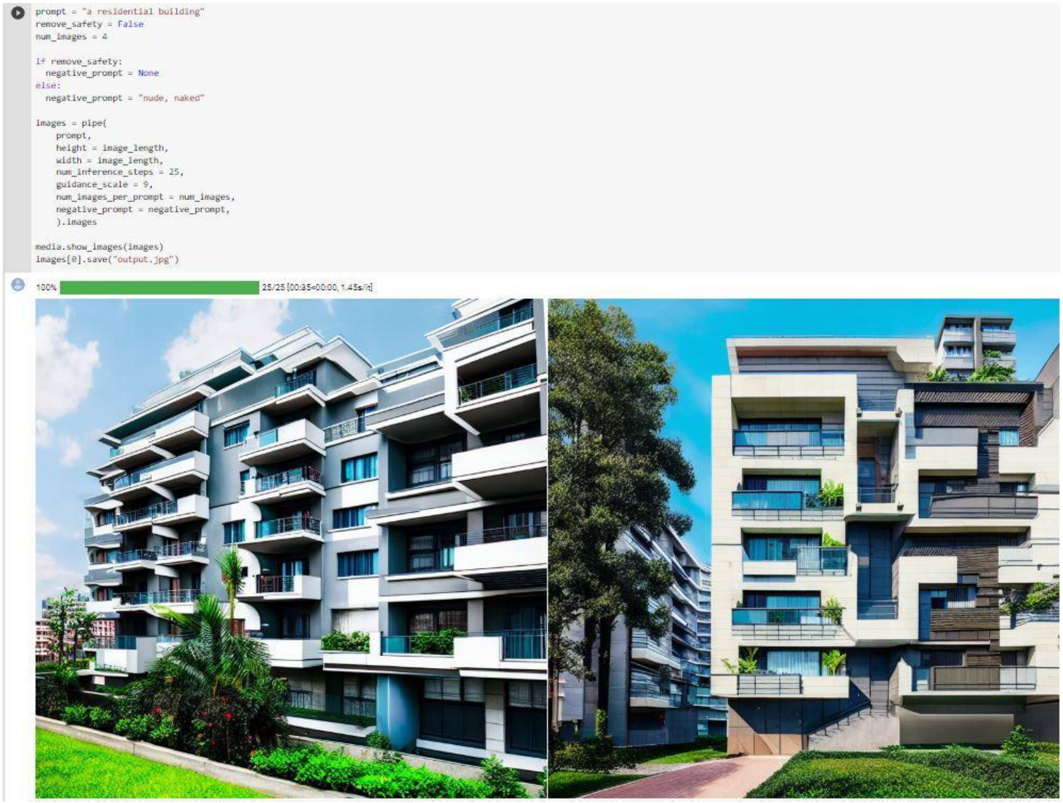


Fig 43: The Stable Diffusion’s google colab used for the experiment.

5.1.1 Design

In this step, we start experimenting with the Stable Diffusion for the general architectural terms that can be used mainly in the design phase

5.1.1.1 Architecture styles

Any architect mostly starts their design based on an initial design style. That's how architecture students are taught at school, to learn from the previous design styles and architects. There are different types of architectural styles that architects use their features in their design. It does not matter if they are very old or super complex to build, existing technologies in design and fabrication make it possible to design different styles. In the following, I test the generator to produce buildings in the most famous styles, from the oldest ones such as Gothic and Baroque, to the more recent ones for example Modern and post-modern. Images for each style are the first five images generated by Stable Diffusion. An example of the prompt is “A building, Gothic style”.

Fig 44: Buildings generated in different architectural styles using Stable Diffusion. Rows from top: Gothic, Renaissance, Baroque, Classical, Art Deco, Modernist, Post-modern, Brutalist, Deconstructivist.

5.1.1.2 Famous Architects' styles

A firm's design style is one of its most distinguishing characteristics. A famous and large firm will inevitably have its design style; however, renowned architects' styles and design thinking are always attractive for architects to consider, especially during the conceptual design and shaping of initial ideas. It's very helpful if AI is capable of differentiating different design approaches and generating designs based on each style. In this step, I asked AI to generate a building in the style of famous architects whose style became famous and is still being used by other designers. The following images are the first five options generated by Stable diffusion. An example of the prompt is "A building, Zaha Hadid style".

Fig 45: Buildings generated in different famous Architects' styles using Stable Diffusion. Rows from top: Antoni Gaudi, Louis Kahn, Le Corbusier, Mis van de Rohe, Peter Eisenman, Frank Gehry, Renzo Piano, Zaha Hadid.

The generated images prove that AI works very well in recognizing different architects' styles and producing new buildings based on their design styles.

5.1.1.3 Building types

In the architecture industry, designing a building can significantly vary based on its functionality and type of it. In this step, I test the capability of AI to produce different types of buildings. We separate the buildings based on their functionality, such as institutional, Cultural, Religious, etc. The following images are the first five options generated by Stable Diffusion. An example of the prompt is “A shopping center building”.

Fig 46: Different commercial buildings generated using Stable Diffusion. Rows from top: office buildings, retail buildings, shopping center buildings.

Fig 47: Industrial buildings generated using Stable Diffusion. Rows from top: Factory buildings, Warehouse buildings.

Fig 48: Cultural buildings generated using Stable Diffusion. Rows from top: library buildings, museum buildings, theater buildings.

Fig 49: Green building generated using Stable Diffusion.

Fig 50: Tower buildings generated using Stable Diffusion. Rows from top: Communication tower buildings, observation tower buildings, skyscraper buildings.

Fig 51: Institutional buildings generated using Stable Diffusion. Rows from top: Hospital buildings, school buildings.

Fig 52: Religious buildings generated using Stable Diffusion. Rows from top: Church buildings, mosque buildings, temple buildings.

Fig 53: House buildings generated using Stable Diffusion. Rows from top: House apartments, house buildings, townhouse buildings.

The variety of generated images based on building functionalities looks very connected and logical.

5.1.1.4 Building forms

As part of this step, I experiment with different building forms. In recent years, the ability to construct complex forms has increased due to technological advances, both in design and fabrication. I believe it is essential to emphasize the importance of building forms in my experiments and to explore how AI can handle them. The experiment focuses on various forms used in a design, such as circular, canted, and triangular. The following images are the first five options generated by Stable Diffusion. An example of the prompt is “A building in the form of a conical”.

Fig 54: Different building forms generated using Stable Diffusion. Rows from top: Canted, circular, conical, faceted, organic, rectangular, spherical, square, triangular.

AI works very well in producing forms relevant to the words used. The only issue with the "Square" form is that the generated images do not represent exactly square buildings in all design options.

5.1.1.5 Building elements

During the conceptual design process, although not much emphasis is placed on the details of buildings, the elements shaping the design of a building, such as doors, windows, roofs, etc., are not considered details, and it is crucial to test whether AI can produce them accurately. In this step, we ask AI to generate the main elements of the building, such as doors, roofs, beams, windows, etc. The following images are the first five options generated by Stable Diffusion. An example of the prompt is “The foundation of a building”.

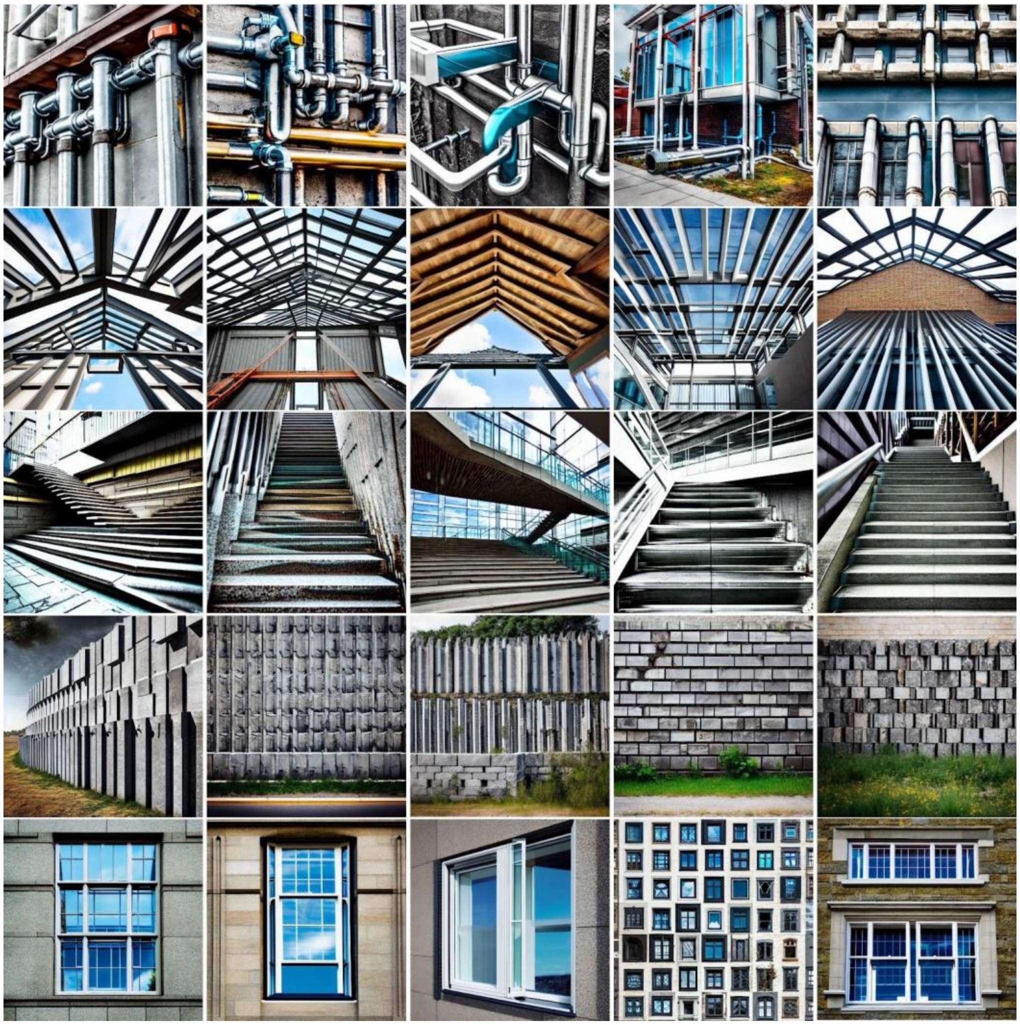


Fig 55: Different building elements generated using Stable Diffusion. Rows from top: Beams, columns, door, electrical systems, elevator, escalator, foundation, plumbing, roof, stairs, wall, windows.

5.1.2 Visualization

In architecture, when designing a project, presenting the idea with amazing visuals has a significant role in any architectural-related project [85]. Design firms spend a lot of money to hire 3D artists to visualize their projects with better quality since it is a key to sell their design. Generative AI has great potential to create stunning computer-generated look-like visuals. This section defines what is essential when visualizing an idea and what features make it look better. We test it with Stable Diffusion to explore its capability in visualization. These visualization factors include landscapes, materials, render styles, etc.

5.1.2.1 Landscape

The landscape includes different scenes that can be added to the design to make it look more realistic. One key feature of the architecture is that buildings must be designed and visualized in different regions and areas. In this step, I experiment with various places, such as streets, parks, forests, etc., with AI to explore how capable it is to position a building in these areas. For this experiment, we test a church building for different scenes. The following images are the first five options generated by Stable Diffusion. An example of the prompt is "A church in the desert".

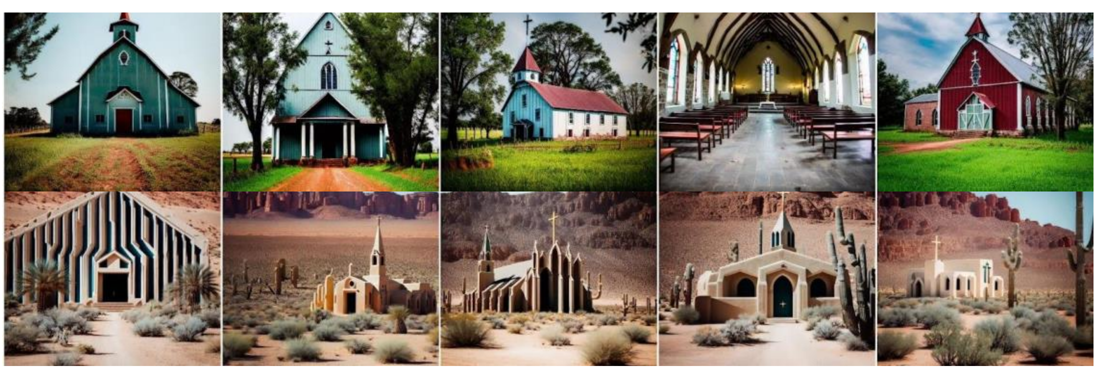


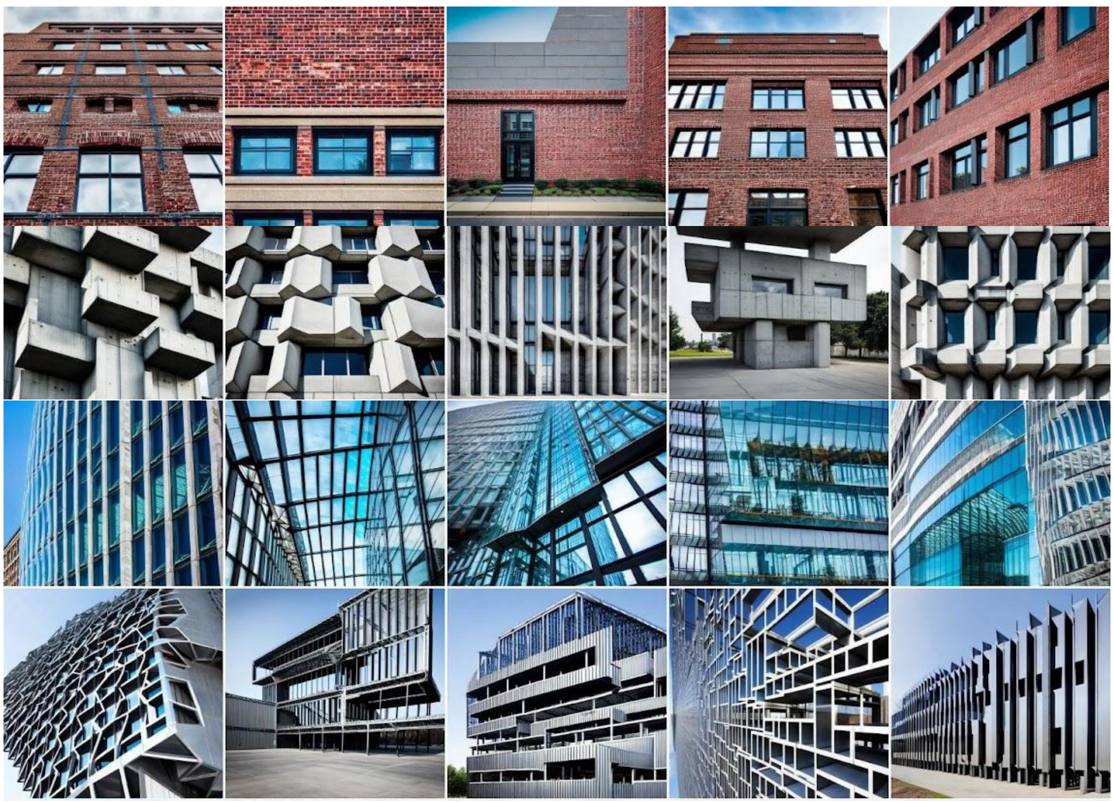
Fig 56: Church building generated in different landscapes using Stable Diffusion. Rows from top: Agricultural, desert, factory, park, beach, forest, lake, mountain, river, waterfall, street.

5.1.2.2 Material, color

In this section, I highlight the importance of color and material in architecture since they have an impact on the appearance of the building. Colors can add contrast and harmony and materials can add pattern and texture to the building. In the first experiment, I tested a door with "Red", "Green," and "Blue" colors. The following images are the first five options generated by Stable Diffusion. An example of the prompt is “A red door”.

Fig 57: Doors generated in different colors using Stable diffusion. Rows from top: Red, green, blue.

The following images are the buildings generated in different materials. I tested them with the most popular and used materials in the architecture industry to make sure AI is capable of understanding all of these different materials. An example of the prompt is “A building with brick material”.



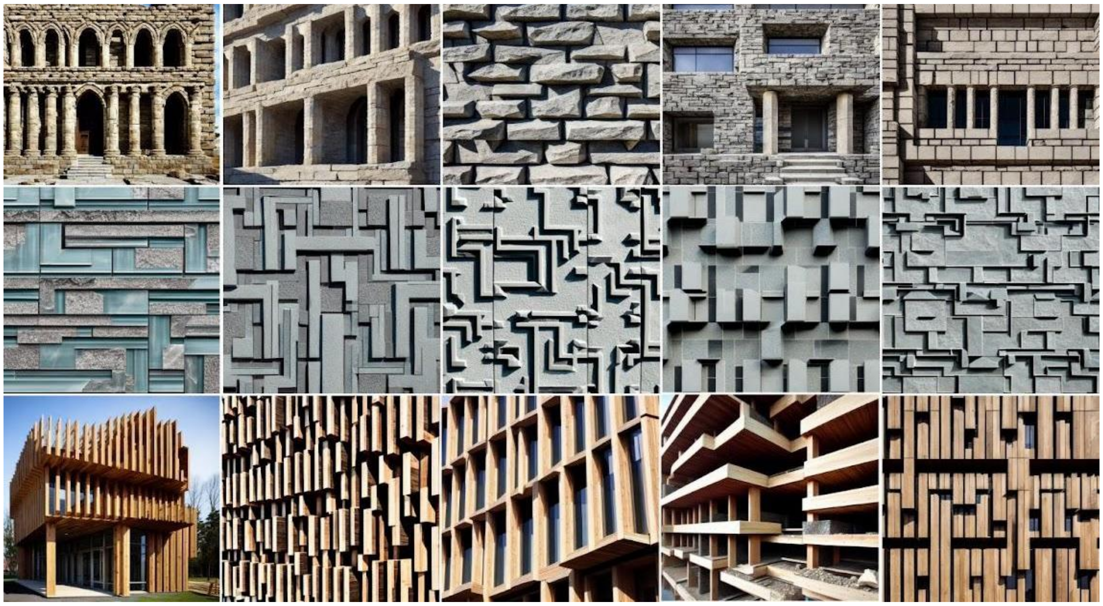


Fig 58: Buildings generated with different materials using Stable Diffusion. Rows from top: Brick, concrete, glass, steel, stone, tiles, wood.

5.1.2.3 Surface types

Another aspect of a building that can be important to consider is the surface type. In this step, I experimented with these features to explore if AI can understand the terms opaque, symmetrical, flat, curved, and others used in architectural design. Controlling them can improve the detail of the visualized design. The following images are the first five options generated by Stable Diffusion. An example of the prompt is "A curved surface".

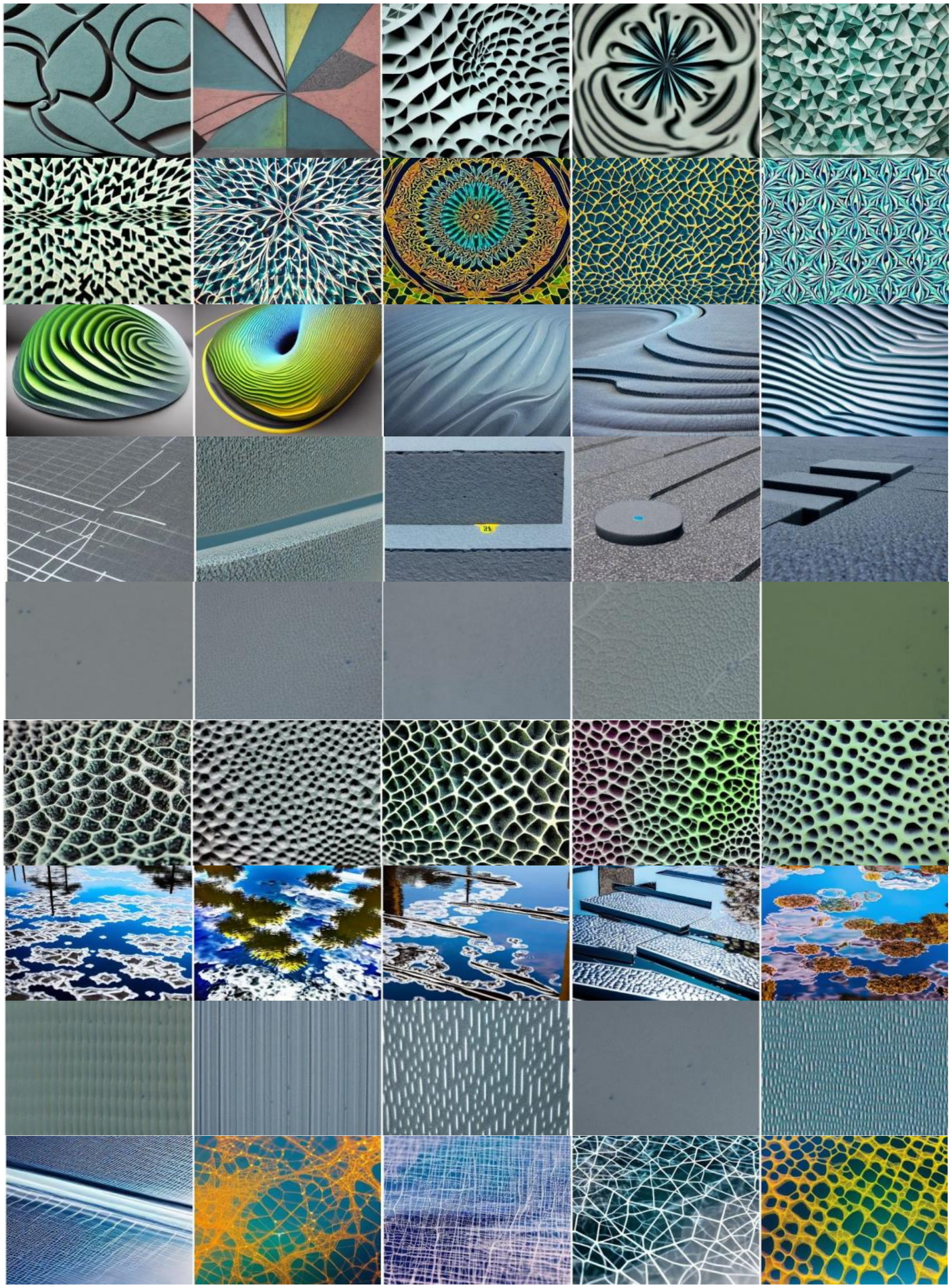


Fig 59: Different surface types generated using Stable Diffusion. Rows from top: Angular, asymmetrical, symmetrical, curved, flat, opaque, organic, reflective, textured, transparent.

5.1.2.4 Architecture views

When preparing architectural documents, it is required to show the project from different views since each presents new information. Other engineers use this information to study the constructibility of the project, which proves its importance. In this experiment, I tested plan, perspective, section, etc. views to determine if AI can distinguish them properly. The following images are the first five options generated by Stable Diffusion. An example of the prompt is “An isometric view of a building”.

Fig 60: Different architecture views generated using Stable Diffusion. Rows from top: Plan view, perspective view, isometric view, elevation view, section view.

AI works pretty well in understanding these architectural views. A few exceptions, such as the section view, also show an elevation in the options. Plan view also shows a bird view as a plan view; however, the error in the generated options is generally insignificant.

5.1.2.5 Render styles

According to the stage of their design, whether in the early stages or the final result, architects and designers prefer to visualize it in different rendering styles. For instance, sketchy styles are used for early-stage ideas, and 3D or photorealistic styles are used for the final results. In this step, we experiment with different rendering styles with AI. The following images are the first five options generated by Stable Diffusion. An example of the prompt “A building, sketch style”.

Fig 61: Buildings generated in different render styles using Stable Diffusion. Rows from top: 3D, artistic, line drawing, photoreal, sketch, watercolor.

5.2 Experiment 2: Comparing Conceptual Designs Using Different Platforms

In this experiment of the thesis, I study Midjorney, DALL-E 2, and Stable Diffusion to generate complex architectural concepts so I can better compare the results. The previous experiment focused on simple but fundamental architectural keywords. In contrast, in this portion, I focused on the more complex concepts that require advanced architectural parametric tools to be designed without the aid of AI.

For the purpose of defining these complex terms, I refer to the most renowned architecture schools and research laboratories pioneering the most advanced design methods. I was inspired to define complex generic concepts by the University College London (Bartlett School of Architecture), Southern California Institute of Architecture (SCI-ARC), and the University of Stuttgart (Integrative Technologies & Architectural Design). Each of these schools explores different approaches to implementing novel technologies in design and fabrication. There is, however, one thing they all have in common: the use of forms from nature to design novel building forms and systems [86]such as the study by O. Krieg, et al., [87] and the development of new solutions for using detailed traditional architectural styles in design by N. Jawad in 2019 [88] and SCI-arc's student projects demonstrate the combination of famous architects' styles to approach novel design solutions [89]. Having studied most of these architecture schools' projects and research over the last few years due to my background in advanced design and digital architecture, I have been able to define complex conceptual design keywords more effectively. Accordingly, I base this experiment on the terminology used in these schools and their research focus and apply it to various building types.

The following diagram illustrates the types of keywords used in the AI experiment to produce complex designs (fig 60).

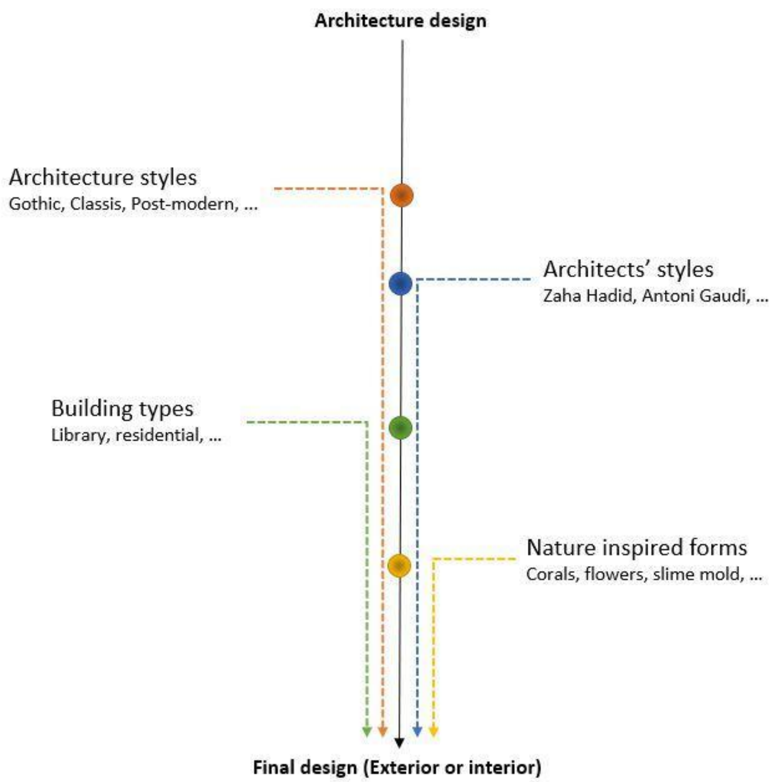


Fig 62: Keywords used for the AI experiments.

Two parts are included in the experiment. Initially, the first part demonstrates the generation of images only based on text. The second part of this experiment employs both visual, and text prompts as inputs.

5.2.1 Text to image

The following images in this experiment show the results of generating designs from the same prompt by Midjourney, DALL-E 2, and Stable Diffusion and testing different text prompts with the same approach.

It is worth noting that for each AI technology, I used the following website and platforms:

Midjourney: Discord

DALL-E 2: OpenAI website

Stable Diffusion: Dream Studio website

As a matter of preference, we would prefer to compare single images per generator, but because Midjourney generates four images by default, we produced more than one design with Stable Diffusion and DALLE-2 to better compare them. To ensure that all tests are created under the same conditions, all the results for each generator are based on the first trial with default settings.

Prompt: A generative tall tower in the desert, brain coral and lavender flower facade, Antoni Gaudi style

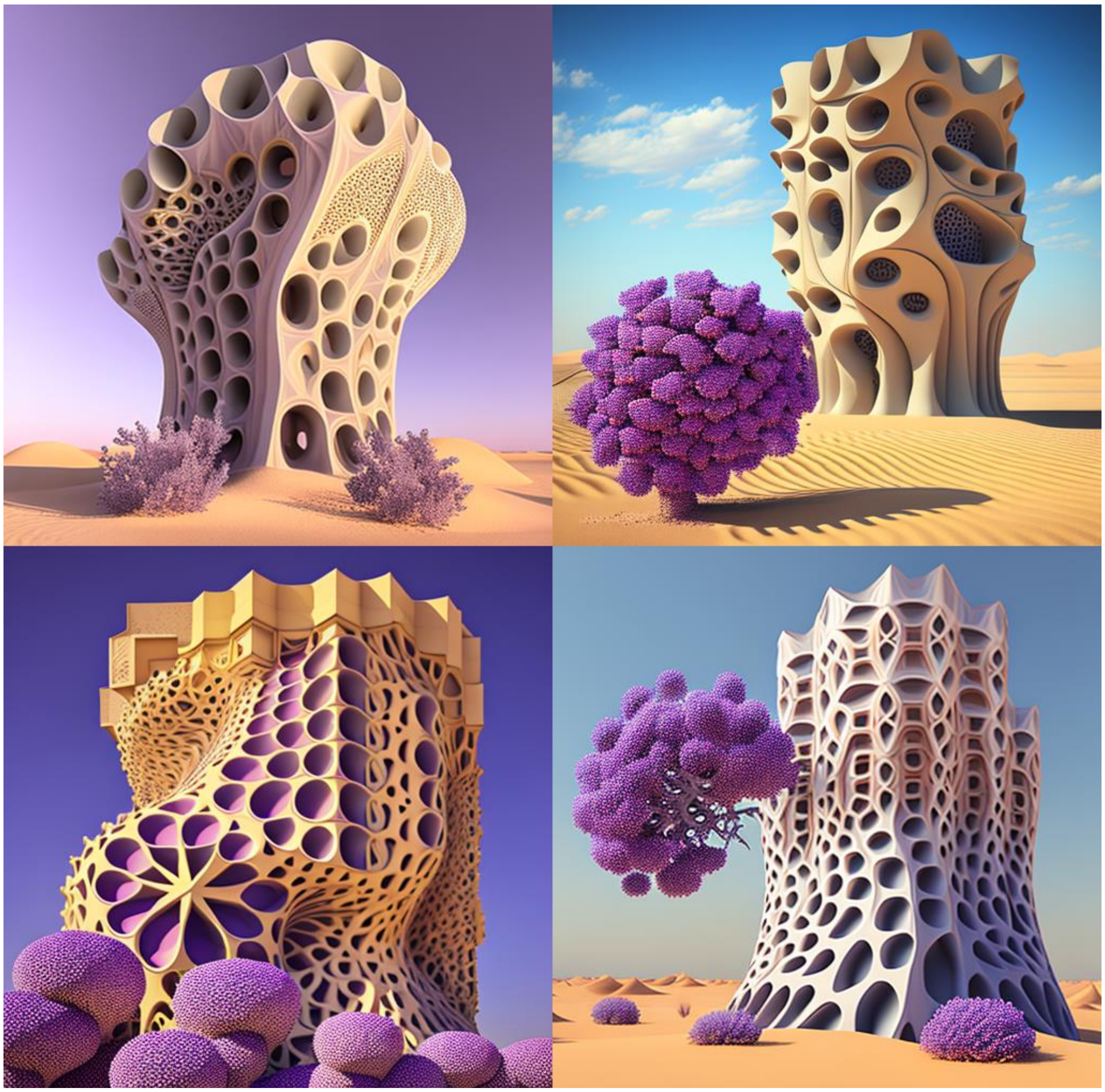


Fig 63: AI generated images of the “A generative tall tower in the desert, brain coral and lavender flower facade, Antoni Gaudi style” prompt by Midjourney. Midjourney perfectly recognized the text and applied it to the design. Integration of the Gaudi style with coral and flower is visible in the result, even though the flower is not fully integrated into the design. The building in the desert, as requested in the prompt.

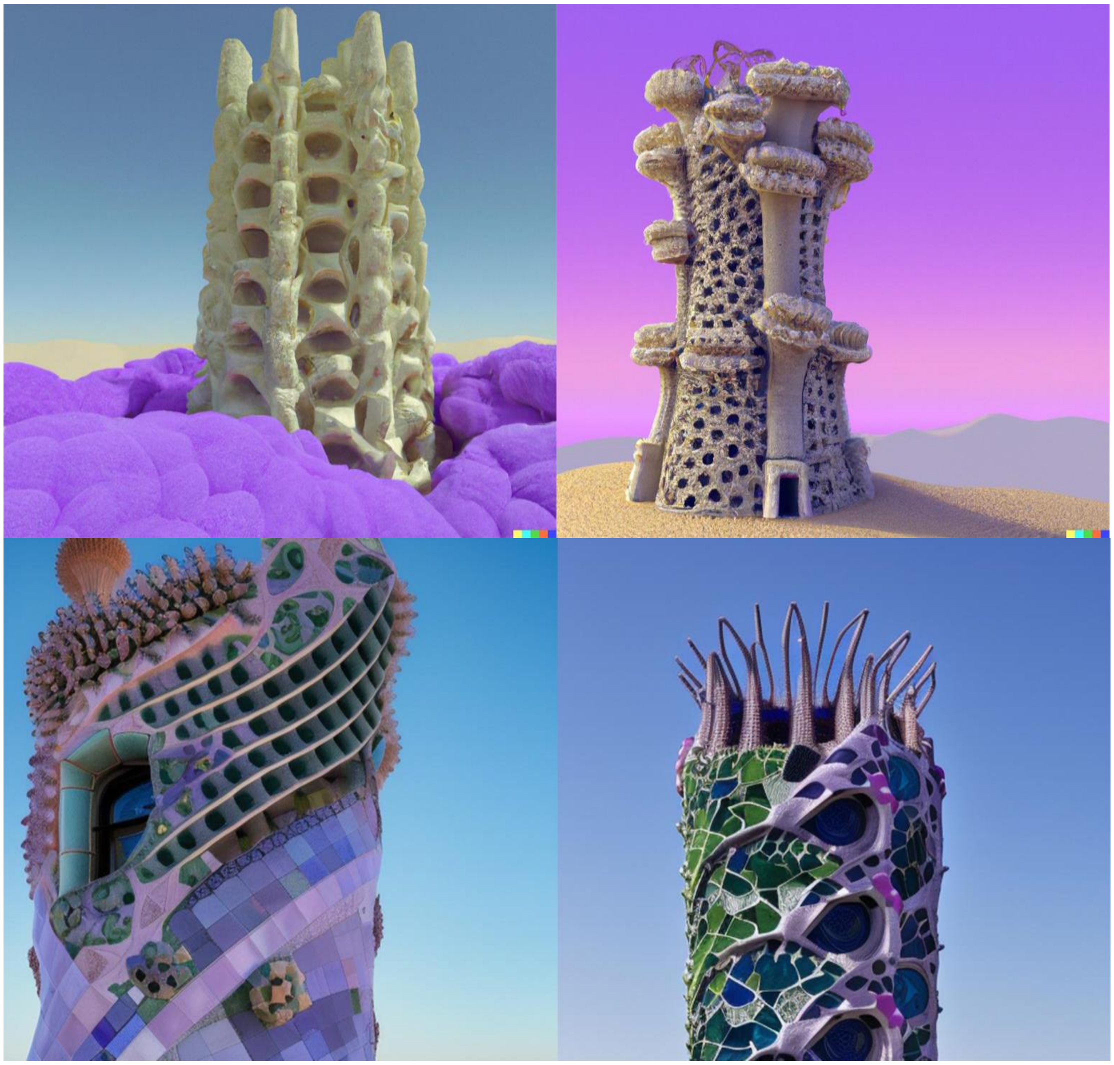


Fig 64: AI generated images of the “A generative tall tower in the desert, brain coral and lavender flower facade, Antoni Gaudi style” prompt. (top) DALL-E 2: The result shows the integration of coral in the tower following the style of Gaudi in the desert or between some flower shape environments; however, the design does not include satisfying details. (Bottom) Stable Diffusion: A beautiful and detailed combination of Gaudi's style with corals and flowers. But what is lacking is the desert environment, as it's not shown in the result.

Prompt: A skyscraper building in the street, lace coral facade

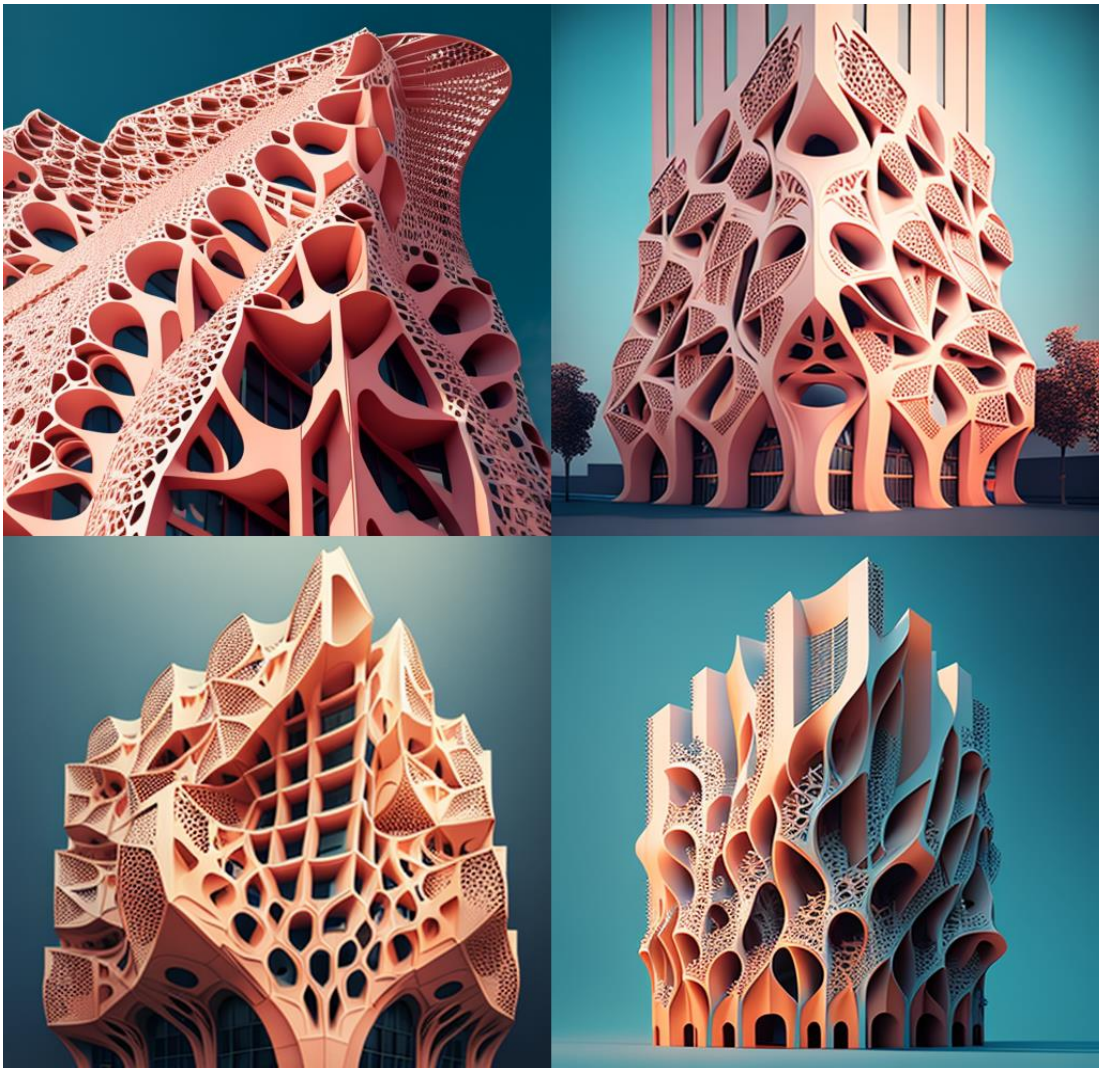


Fig 65: AI generated images of the "A skyscraper building in the street, lace coral facade" prompt. Midjourney: Lace coral is nicely implemented in the design, but the entire building and street are not visible in all the results.

Fig 66: AI generated images of the “A skyscraper building in the street, lace coral facade” prompt. (top) DALL-E 2: The implementation is far from the text prompt, except for the type of building. (bottom) Stable Diffusion: It produces exciting coral forms for the facade. The highlight is that it generated the coral facade as a second layer on the building. The street is lacking, which is not shown in the result.

Prompt: A museum building in the park, coral design, organic style

Fig 67: AI generated images of the "A museum building in the park, coral design, organic style" prompt. Midjourney: The type of generated building could be represented as a museum; moreover, coral and organic forms are nicely implemented in the design. It also includes the park in the design.

Fig 68: AI generated images of the “A museum building in the park, coral design, organic style” prompt. (top) DALL-E 2: It generated a super realistic result with a park background. The coral design is also implemented in the design, but the building type is not very clear. (bottom) Stable Diffusion: In one of the results, the building is nicely designed in the park as a museum in a coral form. The second result, even though it does not look like a building, generates organic forms.

Prompt: A museum building in the street, lace coral and iris flower facade

Fig 69: AI generated images of the "A museum building in the street, lace coral and iris flower facade" prompt. Midjourney: Coral and flowers are nicely integrated into the design as a facade. The type of the building and location looks correct.

Fig 70: AI generated images of the “A museum building in the street, lace coral and iris flower facade” prompt. (top) DALL-E 2: Coral is not visible in the design. The type of building and location looks correct. (bottom) Stable Diffusion: It does a great job of combining coral and flower as a second layer (facade) on a museum building and locating it in the street.

Prompt: A house building in the forest, coral design, Zaha Hadid style

Fig 71: AI generated images of the "A house building in the forest, coral design, Zaha Hadid style" prompt. Midjourney: The style of Zaha Hadid has integrated adequately with coral and is located in the forest. The type of building could be a house and looks correct.

Fig 72: AI generated images of the "A house building in the forest, coral design, Zaha Hadid style" prompt. (top) DALL-E 2: A super realistic integration of Zaha Hadid style in a coral form as a house in the forest. All the texts are visible in the result. (bottom) Stable Diffusion: All the words from the prompt are visible in the result.

Prompt: A mosque building in the mountain, colorful material, Frank Gehry and Zaha Hadid style

Fig 73: AI generated images of the "A mosque building in the mountain, colorful material, Frank Gehry and Zaha Hadid style" prompt. Midjourney: The combination of the two styles is visible in the design. The type of the building is far from the traditional approaches to designing a mosque, which makes it attractive. Colorful material and a mountain environment are included in the design.

Fig 74: AI generated images of the “A mosque building in the mountain, colorful material, Frank Gehry and Zaha Hadid style” prompt. (top) DALL-E 2: The style of Gehry is more visible in the design compared to Hadid. The colorful material is nicely and logically done. (bottom) Stable Diffusion: The results look attractive from architecture point of view; however, the type of building on one of the results could not be a mosque. Also, the mountain is not visible in one of the results.

Prompt: A church building in the river, Antoni Gaudi and Zaha Hadid style, colorful

Fig 75: AI generated images of the “A church building in the river, Antoni Gaudi and Zaha Hadid style, colorful” prompt. Midjourney: The design properly follows Gaudi and Hadid's styles, but the colorful word is applied to the entire results style. River is nicely added to the environment.

Fig 76: AI generated images of the "A church building in the river, Antoni Gaudi and Zaha Hadid style, colorful" prompt. (top) DALL-E 2: The style and type of the building are correctly applied. A colorful material is nicely applied to the building. (bottom) Stable Diffusion: Attractive forms, following the building type and combination of styles, next to the river.

Prompt: A library building in the forest, slime mold design, Antoni Gaudi style

Fig 77: AI generated images of the "A library building in the forest, slime mold design, Antoni Gaudi style" prompt. Midjourney: The style of Gaudi and slime mold are nicely integrated. The building looks close to being a library and is located in the forest.

Fig 78: AI generated images of the “A library building in the forest, slime mold design, Antoni Gaudi style” prompt. (top) DALL-E 2: The results only show Gaudi's style, and nothing can be found as slime mold. The building type can be considered a library. (bottom) Stable Diffusion: Interesting combination of Gaudi's style with slime mold in the forest.

Prompt: An interior of a gothic church building, soft coral design, brick material, Antoni Gaudi style, people inside the building

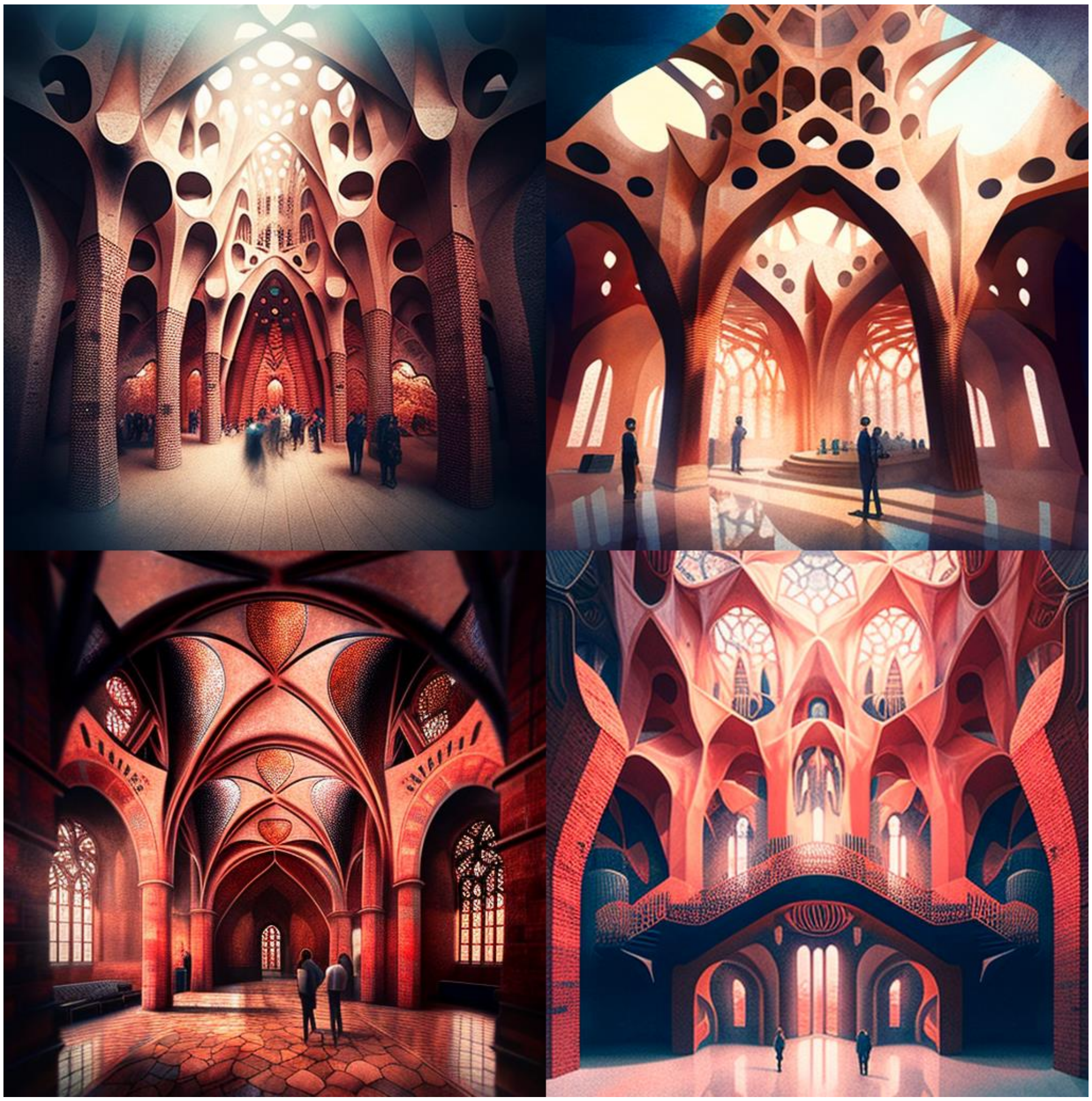


Fig 79: AI generated images of the "An interior of a gothic church building, soft coral design, brick material, Antoni Gaudi style, people inside the building" prompt. Midjourney: An excellent generation of a gothic church in the style of Gaudi using bricks respecting details. However, soft corals are hard to detect. People are correctly located in the building, with the correct scale.

Fig 80: AI generated images of the “An interior of a gothic church building, soft coral design, brick material, Antoni Gaudi style, people inside the building” prompt. (top) DALL-E 2: The results are relatively representing the style of Gothic and Gaudi. Soft corals can be seen in one of the designs. The brick material is not very clear, but people are correctly located in the building. (bottom) Stable Diffusion: The results can be considered Gothic and Gaudi style; however, brick material and people cannot be found.

Prompt: An interior of a library building, coral walls design, Antoni Gaudi and Zaha Hadid style

Fig 81: AI generated images of the “An interior of a library building, coral walls design, Antoni Gaudi and Zaha Hadid style” prompt. Midjourney: Coral is well integrated into both Gaudi's and Hadid's styles. The building looks reasonably like a library.

Fig 82: AI generated images of the “An interior of a library building, coral walls design, Antoni Gaudi and Zaha Hadid style” prompt. (top) DALL-E 2: The space looks like a library, and the style is closer to the Zaha Hadid. Coral design is visible as well. (bottom) Stable Diffusion: The walls are perfectly designed with corals and integrated very well with the interior of a library. The design style is strongly visible.

Prompt: An interior of a living room, Art Deco organic style, slime mold iris flower design furniture

Fig 83: AI generated images of the "An interior of a living room, Art Deco organic style, slime mold iris flower design furniture" prompt. Midjourney: Unique forms are generated by slime mold and iris flower. Art Deco's style is visible.

Fig 84: AI generated images of the “An interior of a living room, Art Deco organic style, slime mold iris flower design furniture” prompt. (top) DALL-E 2: Slime mold is not used in the design, but there are a few flowers. It's hard to see the style in the results. (bottom) Stable Diffusion: Slime mold and flowers are interestingly integrated with the living room and furniture design.

Prompt: An interior of a living room, Baroque organic style, slime mold iris flower design furniture

Fig 85: AI generated images of the “An interior of a living room, Baroque organic style, slime mold iris flower design furniture” prompt. Midjourney: Slime mold and flowers are integrated pretty well into furniture and wall. Baroque style is also visible.

Fig 86: AI generated images of the “An interior of a living room, Baroque organic style, slime mold iris flower design furniture” prompt. (top) DALL-E 2: Only furniture could be represented as a style of baroque. Flowers can be seen in the results. (bottom) Stable Diffusion: Beautiful Baroque living room design; however, slime mold and flowers are not well defined and integrated.

Prompt: An interior of a living room, brutalist organic style, slime mold bubble coral design furniture

Fig 87: AI generated images of the "An interior of a living room, brutalist organic style, slime mold bubble coral design furniture" prompt. Midjourney: Slime mold and bubble corals are designed very well in the living room. Brutal style can also be seen properly.

Fig 88: AI generated images of the “An interior of a living room, brutalist organic style, slime mold bubble coral design furniture” prompt. (top) DALL-E 2: Coral and slime mold are very well integrated into the design. Brutal style is visible in the results. (bottom) Stable Diffusion: The results show the slime mold and bubble coral in the furniture with a brutal style.

Prompt: An interior of a living room, classical organic style, slime mold design furniture

Fig 89: AI generated images of the "An interior of a living room, classical organic style, slime mold design furniture" prompt. Midjourney: Slime mold is mainly applied to the walls or its color to the furniture. The style looks correct.

Fig 90: AI generated images of the “An interior of a living room, classical organic style, slime mold design furniture” prompt. (top) DALL-E 2: A deplorable result in including all the words except the style in one of the results. (bottom) Stable Diffusion: The result hardly follows the style. Slime mold is not visible in the design.

Prompt: An interior of a living room, Gothic organic style, slime mold iris flower design furniture

Fig 91: AI generated images of the "An interior of a living room, Gothic organic style, slime mold iris flower design furniture" prompt. Midjourney: Creative forms generated by slime mold and iris flowers in the style of Gothic. Details are well integrated.

Fig 92: AI generated images of the “An interior of a living room, Gothic organic style, slime mold iris flower design furniture” prompt. (top) DALL-E 2: The result mainly shows the living room and the color of the flowers on the furniture. But the style and slime mold are not visible. (bottom) Stable Diffusion: One of the results is not representing the style and the design elements. The second design generates attractive slime mold furniture with flowers on the ceiling, but the style is not properly included in the design.

Prompt: An interior of a living room, Renaissance organic style, slime mold design furniture

Fig 93: AI generated images of the "An interior of a living room, Renaissance organic style, slime mold design furniture" prompt. Midjourney: Style and slime mold are integrated very well. Slime mold is designed both on the wall and furniture.

Fig 94: AI generated images of the "An interior of a living room, Renaissance organic style, slime mold design furniture" prompt. (top) DALL-E 2: Slime mold is designed nicely in a concrete form following the style. (bottom) Stable Diffusion: There is no organic or slime mold design in the result. Only the renaissance style is visible in the images.

5.2.2 Text + image to image

One of the features of these AI platforms is the ability to use both images and text as input to generate a new result. In other words, to control the image with a text description. This feature is quite helpful in architectural design since it allows designers to generate a variety of design options based on their base image input. In this part of the experiment, I use a simple image of a building and use it as a base for all three generators. The objective is to explore if the generators can understand the base image and generate various design options based on the new text prompt. The following diagram shows the base image used for all experiments in this section (fig 93).

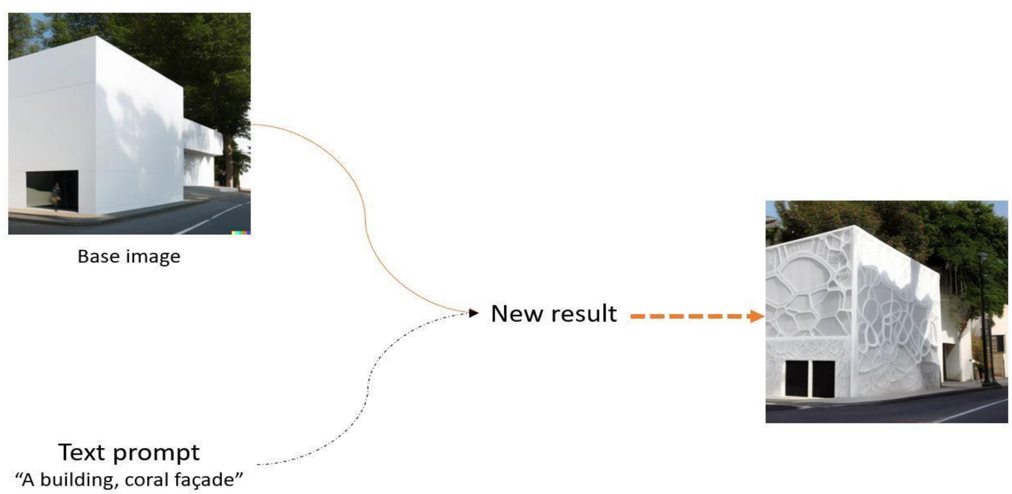


Fig 95: An example of generated design with Stable Diffusion by using both text and image as input.

Prompt: A building, Antoni Gaudi style

Fig 96: AI generated images of the "A building, Anotni Gaudi style" prompt on a base image. Midjourney: It almost keeps the perspective view and trees in the background. The design perfectly shows Gaudi's style.

Fig 97: AI generated images of the “A building, Antoni Gaudi style” prompt on a base image. (top) DALL-E 2: The perspective view looks correct, but the design style is applied only to one of the results. (bottom) Stable Diffusion: View and design options match the prompt and input image entirely. Design options are aligned with Gaudi’s style.

Prompt: A building, classical style

Fig 98: AI generated images of the "A building, classical style" prompt on a base image. Midjourney: The view and style match; however, it does not use the input image precisely for the final design.

Fig 99: AI generated images of the “A building, classical style” prompt on a base image. (top) DALL-E 2: The view is correct, but the style is not appropriately implemented in one of the results. (bottom) Stable Diffusion: The view and styles are correct, but the results are not representing new ideas.

Prompt: A building made of concrete mesh, Frank Gehry style

Fig 100: AI generated images of the “A building made of concrete mesh, Frank Gehry style” prompt on a base image. Midjourney: The design options are very creative and follow Gehry's style. The material and views look correct, but the mesh is not visible in all results.

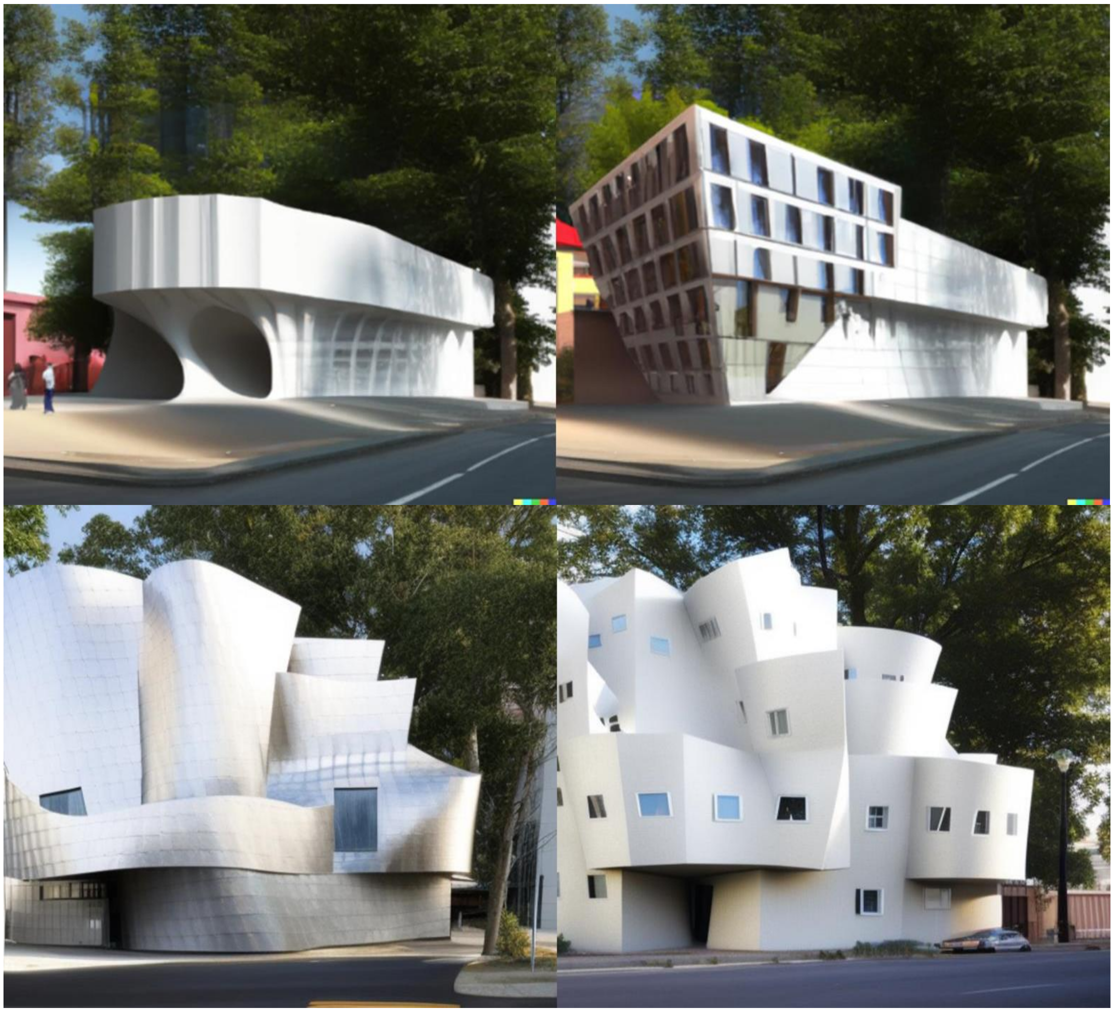


Fig 101: AI generated images of the “A building made of concrete mesh, Frank Gehry style” prompt on a base image. (top) DALL-E 2: The view and material look correct. The design style is also correctly applied but could have better details. (bottom) Stable Diffusion: Outstanding design proposals following Gehry's style: The material in one of the results looks more like metal than concrete.

Prompt: A building, gothic style

Fig 102: AI generated images of the “A building, gothic style” prompt on a base image. Midjourney: Gothic style is integrated correctly with the base image, resulting in exciting designs.

Fig 103: AI generated images of the “A building, gothic style” prompt on a base image. (top) DALL-E 2: The Gothic style is visible in the new results following the same view as the input image. (bottom) Stable Diffusion: The view perspective is quite similar to the base image. Gothic style is also visible but not representing new ideas.

Prompt: A building made of lace coral facade

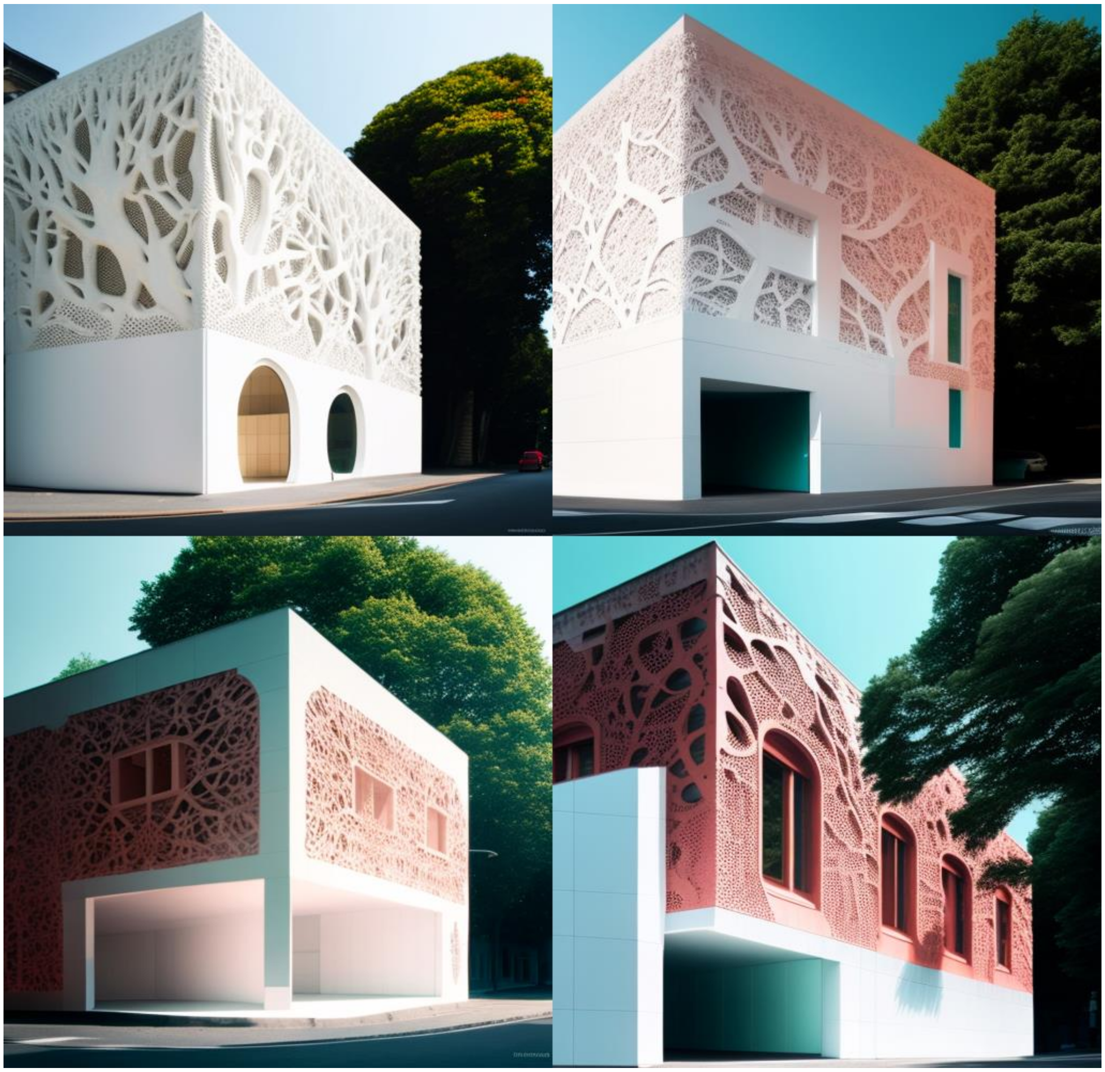


Fig 104: AI generated images of the "A building made of lace coral facade" prompt on a base image. Midjourney: Unique integration of the lace coral with the base image. Smart different material choices.

Fig 105: AI generated images of the “A building made of lace coral facade” prompt on a base image. (top) DALL-E 2: Unsatisfactory results of combining corals and the base image. The design options could be more detailed, attractive, and logical. (bottom) Stable Diffusion: The results are unique and interesting. The base image is properly used in a correct view, and corals are nicely added to the building.

6 Conclusion

In this thesis, I explore and present how architects and designers can make use of generative AI tools (Midjourney, DALL-E 2, and Stable Diffusion) in creating novel conceptual designs. As a first step, I surveyed existing architecture generative tools that can be compared with these AI generators in terms of their ability to produce complex designs. Secondly, since this thesis focuses on generative AI, I examined machine learning architectures used in image-based generative AI and presented comprehensive examples of how the most popular AI tools (Midjourney, DALL-E 2, and Stable Diffusion) are able to produce novel images based on speculative ideas.

I present existing architectural generative tools in order to demonstrate that, in architecture, these tools are used when dealing with complex concepts; however, the new artificial intelligence tools are capable of handling complex concepts and being used early in the design process. Thus, as an experienced designer who has used both generative architecture software and generative AI, I expressed my own opinion regarding the application of AI in conceptual design. It helps better understand the advantages of these AI tools in design. Furthermore, I reviewed a few examples of how AI can be applied to architecture based on the perspectives of famous architects.

It has been mentioned previously that architects randomly experiment with these generative AI tools; however, the potential of each generator to better understand designers' input to integrate all input into output has yet to be surveyed and explored.

Thus, in the last chapter of this thesis, I have used these AI tools to produce the results from the same text prompt to compare their effectiveness in producing architectural design concepts. The experiments are divided into two main categories. For the first step, I explored if AI is capable of producing images based on architectural terminologies. In other words, if AI can properly understand architecture keywords and produce accurate results. I tested different words used in design and visualization based on resources and the most used generic terms according to my experience in design and architectural rendering. For this test, I used Stable Diffusion due to its open-source nature and the fact that there are fewer limitations for a specific number of trials.

The second part of this experiment examined how Midjourney, DALLE-2, and Stable Diffusion can produce complex designs that are hard for designers and existing architecture software to model. A design can either be created from text alone or from both text and images combined. the objective is to show each generator's ability to produce designs with a minimal amount of effort in writing a prompt. Essentially, the goal is to identify which generator produces the most accurate results using simple texts written by anyone who is not an expert in prompt engineering. The words used for this test are derived from the pioneering architecture schools' research covering novel subjects. These keywords include different architectural styles, the styles of famous architects, as well as complex forms found in nature.

All three of these platforms provide powerful tools for generating design concepts, and architects or anyone interested in using them can utilize any of these tools for their design needs. Nevertheless, this part only expresses my observations from these tests. In experiments using only text to generate designs, I found that Midjourney could recognize all the words in the prompt and use them in the generated images. As far as design is concerned, Midjourney can also produce novel designs that follow the main idea of the prompt. In the case of DALLE-2, not all of the words from the prompt were successfully incorporated into the design. However, it is capable of producing realistic-looking designs. Stable Diffusion, on the other hand, is effective in producing innovative designs. While the results of Stable Diffusion do not include all the words from the prompt, and the images do not produce images with good quality and detail, they are acceptable for conceptual design purposes.

During the experiments that involved both text and images, based on my observations, Midjourney produced an array of intriguing results, but it did not retain the reference image in terms of the position and view of the object. It remains a great option to use Midjourney for this task due to its novel results. Despite the fact that DALLE-2 retains its reference image as the basis for the results, the designs lack detail and novelty. All of the texts were not included in the final design of DALLE-2. Stable Diffusion appears to be an ideal tool for this task, considering that it maintains the reference image while producing a number of novel designs on top of it. Stable Diffusion and Midjourney are both suitable options for this approach; however, Stable Diffusion's ability to position the target object and propose design over it makes it an ideal tool from my perspective.

The results indicate that each generator may have different capabilities based on the preferences and needs of different individuals. According to my experience in architectural design, the results are adequate for competing with existing architectural generative tools for the conceptual design stage. Design proposals produced by these AI tools are rapid, they can deal with complex concepts and propose logical architectural design solutions, and their detail, such as material and lighting, are important factors that enable us to consider these AI tools for architectural design.

Future experiments could include studying AI in creating 3D content in architecture. In this thesis, I studied the text-to-image system in conceptual design, but the advancement of generative AI and the ability to generate 3D geometries from a text can revolutionize the architecture industry. It assists in integrating the 3D proposal by AI into the more advanced stages of the design process. In other words, AI-generated 3D models can be used for more serious analysis to facilitate further design development.

Fig 106: AI generated image of the “Shocked architects looking at the AI generated organic structure physical models on the table” prompt, using Midjourney. Source: Created by the author.

Bibliography

[1] Lim. Keunuk, Lee. Hyuk-Jin, "A Study on Works of Antoni Gaudi in Barcelona as UNESCO World Heritage," *Journal of the Association of Korean Photo-Geographers, vol. 21*, 2011.

[2] Eisenman. Peter D, The Formal Basis of Modern Architecture, L. Müller, 2006., Original Submitted in August 1963.

[3] Ragheb, J. Fiona, et al, Architect, Guggenheim Museum Publications, 2001.

[4] Priceonomics, "The Software Behind Frank Gehry's Geometrically Complex Architecture," 2015. [Online]. Available: https://priceonomics.com/the-software-behind-frank-gehrys-geometrically.

[5] Gehry. Frank, et al, "Empowering Design: Gehry Partners, Gehry Technologies and Architect-Led Industry Change," *Architectural Design, vol. 90,* 2020.

[6] R. Garcia Alvarado, "The Control Of Shape: Origins Of Parametric Design In Architecture In Xenakis, Gehry And Grimshaw," *METU JOURNAL OF THE FACULTY OF ARCHITECTURE,* 2012.

[7] Hadid. Zaha, Betsky. Aaron, The Complete Zaha Hadid, 2009.

[8] Wang. Jie, et al, "Application of BIM in Tunnel Design with Compaction Pile Reinforced Foundation Carrying Carbon Assessment Based on Advanced Dynamo Visual Programming: A Case Study in China," *Sustainability, vol. 14,* 2022.

[9] Smakaev. Ravil Maratovich, Nizina. Tatyana Anatolyevna, "USE OF DYNAMO VISUAL PROGRAMMING ENVIRONMENT WHEN DEVELOPING A BUILDING PROJECT IN AUTODESK REVIT," *Economy, Governance and Lave Basis, no. 2*, 2020.

[10] Subhajit. Das, "Energy Model Machine (EMM) Instant Building Energy Prediction using Machine Learning," *ECAADE*, 2017.

[11] Janssen. Patrick, "Visual Dataflow Modelling - Some Thoughts on Complexity," in *ECAADE*, 2014.

[12] "Future Memory," 2020. [Online]. Available: https://www.futurememoryfoundation.org/.

[13] Bachman. David, 10-parametric-plugins-every-architect-should-know, 2017.

[14] Y. Zhang, "Parametric Method Using Grasshopper for Bulbous Bow Generation," in *2018 International Conference on Computing, Electronics & Communications Engineering (ICCECE), IEEE*, 2018.

[15] Zhang. Yongxing, et al, "Parametric Method Using Grasshopper for Bulbous Bow Generation," in *International Conference on Computing, Electronics & Communications Engineering (ICCECE), IEEE*, 2018.

[16] Menges. Achim, "Material Computation: Higher Integration in Morphogenetic Design," *Architectural Design, vol. 82*, 2012.

[17] Guo. Fei, et al, "Computational Implementation of a Tool for Generative Design of High-rise Residential Building Facades," in *Building Facades, 18th International Conference, CAAD Futures*, 2019.

[18] S. O. Hengel, L. M. G. Feijs, "Exploration of Knitted Spacer Fabrics using Grasshopper," *ArXiv,* 2022.

[19] Johan. Ryan, et al, "BUILDING INTELLIGENCE THROUGH GENERATIVE DESIGN," in *Intelligent & Informed, Proceedings of the 24th International Conference of the Association for Computer-Aided Architectural Design Research in Asia*, 2019.

[20] Oussidi. Achraf, and Elhassouny. Azeddine, "Deep Generative Models: Survey," in *International Conference on Intelligent Systems and Computer Vision (ISCV), IEEE*, 2018.

[21] Gm. Harshvardhan, et al, "A Comprehensive Survey and Analysis of Generative Models in Machine Learning," *Computer Science Review, vol. 38*, 2020.

[22] Bank. Dor, et al, "Autoencoders," *Arxiv,* 2021.

[23] Baldi. Pierre, "Autoencoders, Unsupervised Learning, and Deep Architectures," in *JMLR: Workshop and Conference Proceedings*, 2012.

[24] Asperti. Andrea, et al, "A Survey on Variational Autoencoders from a Green AI Perspective," *SN Computer Science, vol. 2, no. 4*, 2021.

[25] Wei. Ruoqi, Ausif. Mahmood., "Recent Advances in Variational Autoencoders With Representation Learning for Biomedical Informatics: A Survey," *IEEE Access, vol. 9,* 2012.

[26] Girin. Laurent, et al, "Dynamical Variational Autoencoders: A Comprehensive Review," *Arxiv,* 2021.

[27] Gui. Jie, et al, "A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications," *IEEE Transactions on Knowledge and Data Engineering,* 2022.

[28] Wang. Kunfeng, et al, "Generative Adversarial Networks: Introduction and Outlook," *IEEE/CAA Journal of Automatica Sinica, vol. 4, no. 4,* 2017.

[29] Cai. Likun, et al, "Utilizing Amari-Alpha Divergence to Stabilize the Training of Generative Adversarial Networks," *Entropy, vol. 22, no. 4,* 2020.

[30] Karimi-Bidhendi. Saeed, et al, "Fully-automated Deep-learning Segmentation of Pediatric Cardiovascular Magnetic Resonance of Patients with Complex Congenital Heart Diseases," *Journal of Cardiovascular Magnetic Resonance, vol. 22, no. 1,* 2020.

[31] Arjovsky. Martin, "Wasserstein GAN," *arXiv,* 2017.

[32] Chen. Xi, et al, "InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets," *arXiv,* 2016.

[33] Zak. Jost, "InfoGAN-Generative Adversarial Networks Part III," 2017. [Online]. Available: https://www.kdnuggets.com/2017/11/infogan-generative-adversarial-networks-part3.html.

[34] Zhu. Jun-Yan, et al, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," *arXiv,* 2017.

[35] Jay. Sinha, "Train your first CycleGAN for Image to Image Translation," 2021. [Online]. Available: https://blog.jaysinha.me/train-your-first-cyclegan-for-image-to-image-translation/.

[36] Salakhutdinov. Ruslan, Geoffrey E. Hinton, "Deep Boltzmann Machines," in *International Conference on Artificial Intelligence and Statistics*, 2009.

[37] Pang. Lei, Ngo. Chong-Wah, "Mutlimodal Learning with Deep Boltzmann Machine for Emotion Prediction in User Generated Videos," in *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval, ACM*, 2015.

[38] Luo. Calvin, "Understanding Diffusion Models: A Unified Perspective," *arXiv,* 2022.

[39] Yang. Ling, "Diffusion Models: A Comprehensive Survey of Methods and Applications," *arXiv,* 2022.

[40] Ryan. O'Connor, "Introduction to Diffusion Models for Machine Learning," 2022. [Online]. Available: https://www.assemblyai.com/blog/diffusion-models-for-machine-learning-introduction/.

[41] Karagiannakos. Sergios, Adaloglou. Nikolas, "How diffusion models work: the math from scratch," 2022. [Online]. Available: https://theaisummer.com/diffusion-models/?fbclid=IwAR1BIeNHqa3NtC8SL0sKXHATHklJYphNH-8IGNoO3xZhSKM_GYcvrrQgB0o .

[42] Ho. Jonathan, et al, "Denoising Diffusion Probabilistic Models," *arxiv*, 2020.

[43] Weng. Weihao, Zhu. Xin, "INet: Convolutional Networks for Biomedical Image Segmentation," *IEEE Access, vol. 9,* 2021.

[44] Jonathan. Ho, "High Fidelity Image Generation Using Diffusion Models," 2021. [Online]. Available: https://ai.googleblog.com/2021/07/high-fidelity-image-generation-using.html?m=1.

[45] "Google’s new AI can restore low-res photos," 2021. [Online]. Available: https://gadget.co.za/googles-new-ai-can-sharpen-low-res-photos/.

[46] Lugmayr. Andreas, et al, "Repaint: Inpainting using denoising diffusion probabilistic models," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2022.

[47] Meng. Chenlin, et al, "Sdedit: Guided image synthesis and editing with stochastic differential equations," in *International Conference on Learning Representations*, 2021.

[48] Luo. Chitong, Hu. Wei, "Diffusion probabilistic models for 3d point cloud generation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2021.

[49] Nair. Nithin Gopalakrishnan, et al, "IMAGE GENERATION WITH MULTI-MODAL PRIORS USING DENOISING DIFFUSION PROBABILISTIC MODELS," *arXiv*, 2022.

[50] Croitoru. Florinel-Alin, et al, "Diffusion Models in Vision: A Survey," *arXiv*, 2022.

[51] C. Saharia, W. Chan, H. Chang, C. Lee, J. Ho, T. Salimans, D. Fleet, and M. Norouzi, "Palette: Image-to-image diffusion models," in *Proceedings of SIGGRAPH, pp. 1–10*, 2022.

[52] Rombach. Robin, et al, "High-Resolution Image Synthesis with Latent Diffusion Models," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE*, 2022.

[53] "How does Stable Diffusion work?," 2022. [Online]. Available: https://stable-diffusion-art.com/how-stable-diffusion-work/.

[54] J. Alammar, "The Illustrated Stable Diffusion," 2022. [Online]. Available: https://jalammar.github.io/illustrated-stable-diffusion/.

[55] Youssef. Hosni, "Getting Started With Stable Diffusion," 2022. [Online]. Available: https://pub.towardsai.net/getting-started-with-stable-diffusion-f343639e4931.

[56] "Stable Diffusion – AI Art for the Masses," 2022. [Online]. Available: https://tinyurl.com/yrper6e5.

[57] Suraj. Patil, "Stable Diffusion with Diffusers," 2022. [Online]. Available: https://huggingface.co/blog/stable_diffusion.

[58] Gallo. Ignazio, et al, "Semantic Text Encoding for Text Classification Using Convolutional Neural Networks," in *14th IAPR International Conference on Document Analysis and Recognition (ICDAR), IEEE*, 2017.

[59] Li. Saihan, and Bing Gong, "Word Embedding and Text Classification Based on Deep Learning Methods," in *MATEC Web of Conferences, edited by I. Barukčić, vol. 336*, 2021.

[60] Rombach. Robin, et al, "High-Resolution Image Synthesis with Latent Diffusion Models," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE*, 2022.

[61] Ambuj. Pawar, "A First Look at DALL-E 2 — How It Works Under the Hood," 2022. [Online]. Available: https://betterprogramming.pub/dall-e-2-and-why-everyone-is-talking-about-it-baae2e41655e.

[62] Ramesh. Adity, et al, "Hierarchical Text-Conditional Image Generation with CLIP Latents," *arXiv*, 2022.

[63] Alberto. Romero, "DALL·E 2, Explained: The Promise And Limitations Of A Revolutionary AI," 2022. [Online]. Available: https://www.topbots.com/dalle-2-explained/.

[64] Aditya. Singh, "How Does DALL·E 2 Work?," 2022. [Online]. Available: https://medium.com/augmented-startups/how-does-dall-e-2-work-e6d492a2667f.

[65] Ryan. O'Connor, "How DALL-E 2 Actually Works," 2022. [Online]. Available: https://www.assemblyai.com/blog/how-dall-e-2-actually-works/.

[66] "DALLE-2," 2022. [Online]. Available: https://openai.com/dall-e-2/.

[67] Jonas. Oppenlaender, "The Creativity of Text-to-Image Generation," in *25th International Academic Mindtrek Conference, ACM*, 2022.

[68] Lars. Nielsen, "An advanced guide to writing prompts for Midjourney (text-to-image)," 2022. [Online]. Available: https://medium.com/mlearning-ai/an-advanced-guide-to-writing-prompts-for-midjourney-text-to-image-aa12a1e33b6.

[69] Wu. Xintian, et al, "F3A-GAN: Facial Flow for Face Animation With Generative Adversarial Networks," *arXiv*, 2022.

[70] Bardina. M, et al, "Application for Generating Images Using Generative-Adversarial Networks," in *International Multi-Conference on Industrial Engineering and Modern Technologies(FarEastCon), IEEE*, 2020.

[71] David. Newton, "Deep Generative Learning for the Generation and Analysis of Architectural Plans with," *Education and Research in Computer-Aided Architectural Design in Europe and XXIII*, 2019.

[72] Yu. Qiu, et al, "Architectural Facade Recognition and Generation through Generative Adversarial," in *International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE),*, 2020.

[73] Marjanovic. D, et al, "ERROR IN ARCHITECTURAL DESIGN PROCESS: TOWARDS A COGNITIVE MODEL," in *Proceedings DESIGN, the 10th International Design Conference,* , 2008.

[74] Shamsudeen. Musa, "EFFECTS OF DESIGN ERRORS ON CONSTRUCTION PROJECTS," *International Journal of Scientific and Engineering Research,* 2016.

[75] "AI tools like DALL-E 2 and Midjourney are helping architects—and their clients—design new buildings," 2022. [Online]. Available: https://www.fastcompany.com/90780871/ai-tools-like-dall-e-2-and-midjourney-are-helping-architects-an.

[76] DigitalFUTURES, "DigitalFUTURES Talk: The AI Design Revolution: DALLE vs MidJourney," 2022. [Online]. Available: https://digitalfutures.international/talks/.

[77] Poole. Ben, et al, "DreamFusion: Text-to-3D using 2D Diffusion," *arXiv,* 2022.

[78] Nichol. Alex, et al, "Point-E: A System for Generating 3D Point Clouds from Complex Prompts," *arXiv,* 2022.

[79] "VERAS," 2023. [Online]. Available: https://www.evolvelab.io/veras.

[80] OpenAI, "ChatGPT: Optimizing Language Models for Dialogue," 2023. [Online]. Available: https://openai.com/blog/chatgpt/.

[81] parametric-architecture, "Arturo Tedeschi Has Taken ChatGPT, Phyton, And Grasshopper 3D To The Next Level," 2023. [Online]. Available:

https://parametric-architecture.com/arturo-tedeschi-has-taken-chatgpt-phyton-and-grasshopper-3d-to-the-next-level/.

[82] "Veras for Revit 0.8," 2023. [Online]. Available: https://download.archsupply.com/get/download-veras-for-revit/.

[83] F. D. Ching, A Visual Dictionary of Architecture, 1995.

[84] "Architectural Vocabulary," University of North Dakota, 2021.

[85] s. G. Vesselin Gueor, "Architectural Visualization: Understandings And Misunderstandings," in *International Scientific Conference Computer Science,* 2008.

[86] ArchDaily, "Bartlett School of Architecture: The Latest Architecture and News," 2022. [Online]. Available: https://www.archdaily.com/tag/bartlett-school-of-architecture.

[87] O. D. Krieg, "HygroSkin – Meteorosensitive Pavilion," *UCL Press,* 2014.

[88] N. Jawad, "Reviving the Muqarnas in Damascus: a historical architectural form in a contemporary discourse," in *UCL,* 2019.

[89] "Design Studio," [Online]. Available: https://www.sciarc.edu/academics/undergraduate/design-studio.

[90] J. Fontan, "5 Architecture Phases Of Design Explained," 2021. [Online]. Available: https://fontanarchitecture.com/architectural-design-phases/.

[91] S. VERMA, "Architectural Design Process & Its 7 Phases Explained," 2022. [Online]. Available: https://www.oneistox.com/blog/architectural-design-process.

[92] S. Chaillou, "ArchiGAN: a Generative Stack for Apartment Building Design," 2019. [Online]. Available: https://developer.nvidia.com/blog/archigan-generative-stack-apartment-building-design/.