SPATIAL FINDINGS ON CHILEAN ARCHITECTURE STYLEGAN AI GRAPHICS

TOMAS VIVANCO LARRAIN[1], ANTONIA VALENCIA[2] and PHILIP F. YUAN[3]
[1]*Pontifical Catholic University of Chile, Tongji University*
[1]*tvivancol@uc.cl*
[2]*Pontifical Catholic University of Chile*
[2]*arvalencia@uc.cl*
[3]*Tongji University*
[3]*philipyuan007@tongji.edu.cn*

Abstract. The use of StyleGAN algorithms proposes a novel approach in the investigation of architectural images. Even though graphical outcomes produced by StyleGAN algorithms are far from being architectural spaces, they might become a starting point in the creative process of architectural projects. By creating a database of specific categories of architectural images located in certain contexts, significant findings might emerge regarding their categorization in accordance to the 'style of a culture.' This research analyzes the architectural images that result from implementing StyleGAN algorithms in a database of images of Chilean houses built between the years 2010 and 2020 and selected as finalist of the ´Project of the Year´ from international viewers and curators of the most viewed architectural website of the world. Our findings suggest that Chilean houses have two distinctive elements strongly influenced by human bias: the proportion of voids in the architectural-like generative volume and the integration of vegetation or landscape.

Keywords. StyleGAN; Chilean architecture; artificial intelligence; spatial findings.

1. Introduction.

Recent advances in the area of computer science have allowed for the integration of new procedural tools in the design process. Designers and computer scientists have found common ground in exploring and creating new formal possibilities, building a three-dimensional creative space to navigate within its latent vector.

Conceptually, ideas exist due to the fact that we can represent them, and in turn, representation implies the use of known or trained tools or methods that make the idea explicit. At the same time, a new creative idea is driven by intuition, which cannot be understood or described employing logical operations that only become visible when it can surprise an observer or user, generating a certain degree of novelty. Therefore, a new creative idea lies in the recipient of that idea.

PROJECTIONS, Proceedings of the 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2021, Volume 1, 251-260.

From the perspective of the creator, ideas are produced by connections that are not necessarily rational or logical, originated by three sources of information. The first is inherited prior or genetic knowledge; the second is knowledge acquired through feelings and perceptions, or senses; and the third is knowledge gained throughout the creator's life.

These three sources of information nourish our massive database–the place where we explore all the knowledge that feeds into our unique and creative design processes. From this perspective, we then ask ourselves: is it possible to explore the limits of our creativity through the generation of databases of information or images?

Considering the origin of ideas as a starting point for the creative development of design processes questions the idea of creative freedom, as it challenges the notion that an idea is the interconnection of a non-logical organization of senses (Baars, 1988), experiences, learned knowledge and genetics.

On the one hand, computational technologies, specifically Artificial Intelligence, represent a significant advance towards the development, optimization and automation of processes; on the other, there is a large and emerging debate (Boden, 199) about their ability to replace or extend human creativity. Understanding the information that fuels our creativity as "manageable" and "arrangeable" allows for the possibility of generating new images or patterns without direct human intervention (Bechtel & Abrahamsen, 2002) and designing with images or formal references previously known and with which we consciously and unconsciously operate. So, can an algorithm navigate or operate within creativity? Alternatively, perhaps, are algorithms able to explore some of the formal attributes of existing architectural buildings in specific contexts that influence us and define, in an unconscious manner, a local style or attributes?

This research explores and analyzes synthetic images that result from implementing StyleGAN algorithms with a database of images of 1,500 Chilean houses between the years 2010 and 2020 in order to identify spatial attributes that suggest a particular architectural style present in Chilean houses.

1.1. GENERATIVE ADVERSARIAL NETWORKS.

Generative Adversarial Networks (GANs) are generative Artificial Intelligence techniques based on a networked algorithm that recognizes patterns and generates relationships between a large amount of data, similar to the way neural networks work in the brain (Brock, 2019). The input information is processed in layers of algorithms, recognizing patterns and generating associations between them in order to generate a result. If the result is similar to the original, the network learns from it and generates a new result, ensuring that the generated outcome is novel compared to the original database.

GANs are composed of two neuronal networks that 'compete' in order to produce these original images: a bottom-up generator which generates images, and a top-down Discriminator that evaluates the generated images. Somehow, the Generator takes the role of the artist and the Discriminator of the critic (Leach,

2019), as the former has to constantly 'surprise' the latter. In this process of 'competition,' the generator must produce realistic images until the Discriminator is unable to distinguish them from the given dataset.

1.2. GENERATOR AND DISCRIMINATOR STYLEGAN

The Generator and Discriminator are two basic components that enable GAN's functioning. Both components are models of Neuronal Networks that are constantly operating.

The function of the Generator (Figure 1) is to generate fake results from random noise or from an aleatory input given by real data. The Discriminator is the algorithm that distinguishes real data from fake data before re-entering that data to the Generator. This process is called the Epoch.

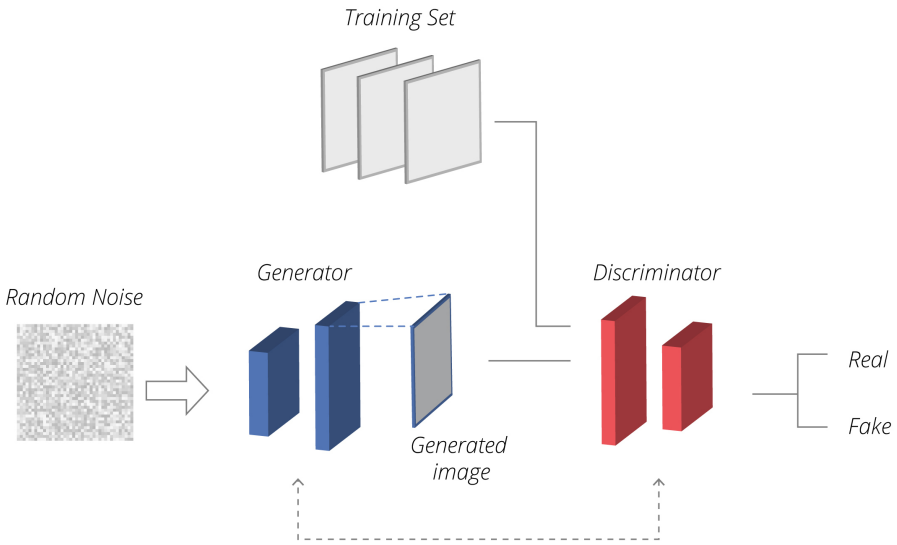

Figure 1. Generator and Discriminator of a Generative Adversarial Network.

1.3. STYLEGAN

Presented by Nvidia in 2018, it occupies the structure of a Generative Antagonist Network (GAN) with the main purpose of generating new images of faces, from a database of images of real faces modified with noise and interference. Two "neural" networks constantly interact, where one proposes new options, and the other evaluates their level of authenticity.

Similar to the process when humans learn a specific architectural style or geometric element (del Campo, Manninger, 2020), neural networks are trained with as many images as possible–at least a thousand–to recognize the differences between elements and styles.

Suppose the process of computing a large number of images to recognize motifs

(Priore, 2001) and to produce new stylish images with those patterns are similar to the human process of observing, classifying and creating. In that case, computers through AI algorithms might become, in a way, as creative as humans or at least become a supportive tool with which to navigate through creativity in an extended and collaborative way.

Human Learning and Machine Learning applied to the analysis of images are similar processes based on the detection of significant patterns. The main difference between these two processes is that AI-based learning is an automatized algorithm (Shalev-Shwartz, 2014) that transforms previous experiences into specific know-hows without knowing what the symbolic, cultural and functional purposes of the analyzed object are. Hence, it cannot make any relationships or functional associations with other objects or generate any results outside the input trained data.

2. Methods and processes.

The human-machine learning design methodology allows the designer to take the role of curator. Co-designing with machine intelligence involves an exploratory territory where the outcome cannot be predicted, and thus involves tweaking specific parameters and objectives to turn a human design process into a human-machine conversation.

The main objective of this research is to generate synthetic architectural images using a StyleGAN algorithm and analyze those images to detect main objects or elements that emerge as a suggestion of style or attributes. To achieve this objective, StyleGAN2 and Object Detection algorithms from Runway ML app were trained to generate the images and to detect objects and elements from them. Four steps composed the process (Figure 2) to accomplish the results and findings:

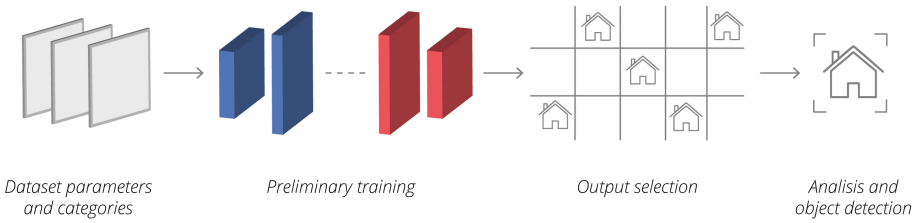

Figure 2. Four steps methodology used on the presented research.

RunwayML is an Artificial Intelligence tool for creators that allows them to discover, create, and use artificial intelligence in creative work. Creativity is being subjected to a revolution with human and machine interactive co-design by allowing creators to use this new tool and explore the boundaries of creation. Designers with no coding experience may take into consideration new design questions and objectives, encouraging experimentation with pre-trained models.

The images were chosen from ArchDaily´s ´Building of the Year´ database from the category of Chilean houses build between 2010 and 2020, each project was selected by their community- with round 160 million views per month- and

global curators composed by architects and designers. This means, there is a crossed, wide and blind selection of each house without any intervention from the authors, giving a global and local validation of their architectural attributes.

2.1. DATASET PARAMETERS AND CATEGORIES

The first step to train a custom model for the generation of StyleGAN based images is to gather a dataset that will serve as the input to train the model. It is crucial to elaborate an organized, clean and robust database and later pre-process it. These steps consist of a collection of around 1,500 images (Figure 3), ensuring that they are representative and exemplify the target image, and that undesired elements are removed.

The database is classified in categories and subcategories to unify it and make it more effective. Considering the role of the designer in the human-machine learning process, the realization of a robust and organized dataset will explain a work strategy, a database classification topology, and the parameters needed to achieve the results.

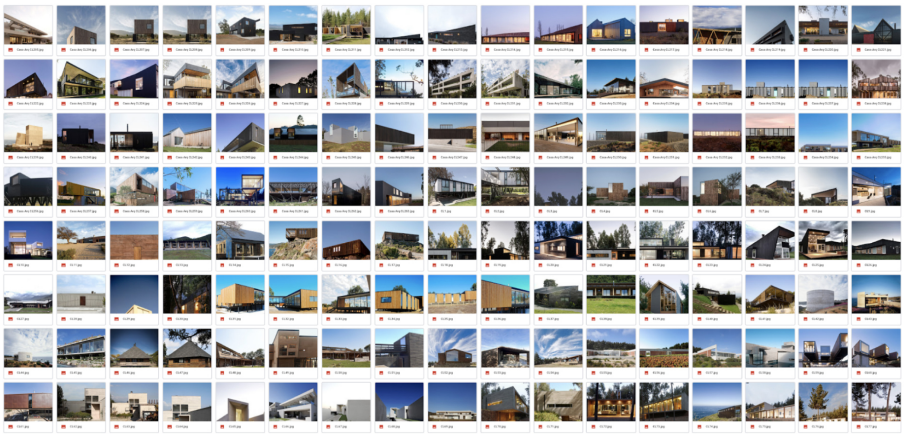

Figure 3. Dataset topology of 1.500 Architectural Chilean Houses facades retrieved from ArchDaily database.

The images were selected from Archdaily's database; each are photographs of the facade of houses built in different regions in Chile. Renowned local architects designed each house, and curators from Archdaily selected them and published them on their platforms. Many of them were published in different national and international architectural magazines. The main characteristic for the selection of images consists of generating a diverse but pre-worked dataset from a diverse range of designs, including houses in urban, rural and seaside contexts, as well as two volumetric strategies: boxed or 'Mediterranean' volume with a minimum roof slope, and triangular volume or Colonial-Chilean house with a steep roof slope. Each image showed a centred facade, showing the surroundings of the house, meaning the image needed to show where the house was built; also, it must show the entire facade, and the volumes needed to be fully visible in the perspective of

elevation views.

2.2. PRELIMINARY TRAINING

The first training of the StyleGAN model was carried out after obtaining a robust and pre-worked database. First, it began with the choice of a pre-trained model to generate faces from the Flickr database, to then set the number of training steps and refine specifications. Once the pre-processing was ready, a latent space allowed the navigation from an extensive arrangement of 512 dimension vectors (Figure 4), each one of them offering infinite possibilities of synthetic images. The possibilities are endless, and curatorship takes a critical role in evaluating the outcomes.

Figure 4. A generative vector space grid as a result of the training process.

The algorithm was trained under 2,000 steps, each one of them understood as an epoch. As can be seen in Figure 4, the pre-trained model 'forgot' about human faces and only selected facades, ensuring a high-performance result and a generation of a clean image.

Truncation and sampling distance mark the variability of the output images within the 512 dimension vectors. An increment in the sampling distance generates a more diverse set of facades, resulting in a more significant distinction between picture and picture in the latent space. Truncation, on the other hand, determines how "realistic" the output image will be. The closer the number is to 0, the closer it approximates to the original dataset (Table 1). For example, using an input vector of an image as the centre of the starting point (of the latent space), and lowering down the truncation-psi and the sampling distance, results in the effect of detail variation under a similar input shape.

Table 1. Variation of truncation and sampling distance with an input vector as a starting point in the latent space. Images represent the difference between each selected stop in the parameters and the effect on the generated images.

2.3. OUTPUT SELECTION

As previously stated, the database is a critical piece that feeds the system. Using clear and concise input criteria will provide thorough work for the outcome images. In some way, the input criteria are the borders where the generated images cannot pass. The cognitive process of selecting input images and evaluating the outcomes based on specific classifications identified in the generative process means an unconscious but culturally constructed bias that controls the process.

Why does one image pass the filter, and another does not? What does the designer consider when selecting an image that represents the style of Chilean architectural facades? The act of analysis and selection is essential, as those parameters affect the meta-designer/curator level and final selection.

The vector latent space images were selected under three objective criteria based on the dataset topology: first, the clarity of the image, where the edges of the house-like volumes were clean and straight, this because all the images from the datasets showed volumes straight lines; second, clear boxed or 'Mediterranean' volumes; and third, Chilean colonial house type volumes with steep roof slopes.

To the human eye, each selected outcome image presented some evident characteristic that defined standard classification criteria, suggesting the definition of a particular "style" of a Chilean house over others, from which, after analyzing the dataset and the generated images, two main criteria were observed:

(1) Incidence of vegetation in the image as an architectural element.

(2) Relationship between voids and solids of the images, with a strong tendency to voids.

2.4. ANALYSIS AND OBJECT DETECTION

From the latent vector space, 1,000 synthetic images were selected after finishing the first training. The selection of each image depended on the clarity of the architectural volume, vegetation as an external but integrated element, and voids inside the architectural volume (Figure 5).

With those three elements, an Object Detection algorithm (pre-trained to detect different objects or elements in images) was firstly trained with the original dataset to recognize three types of elements: vegetation, windows or façade voids, and volumes (Figure 6). A second training was done using the vectors of the generated synthetics images, and, at last, each synthetic image was analyzed to recognize and meet the previously established criteria.

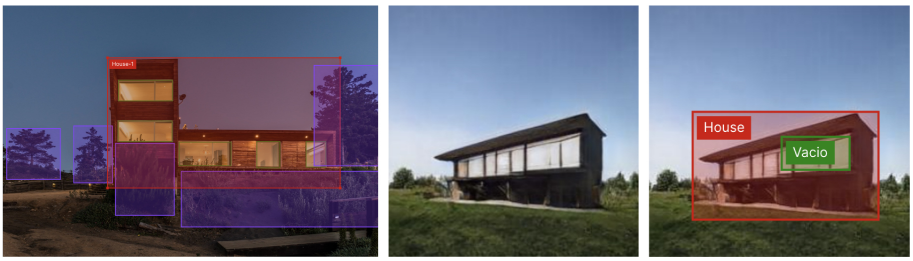

Figure 5. Object detection model trained to identify the classified elements that compose a Chilean architectural facade: presence of vegetation, and proportion of empty v/s full and identify a house.

Figure 6. Selected images by the object detection algorithm based on criteria.

3. Final training images and results.

The results of the StyleGAN2 algorithm are images that look like houses, even though they were just the combination of pixels with supposedly no spatial relationship.

Nevertheless, two main spatial standard features that emerged from the process are the integration of landscape and nature, and the proportion of the volume-void relationship. Both characteristics serve to analyze the generated images and to define a specific Chilean Houses Architectural Style characteristics (Figure 7).

Figure 7. Final Synthetic images after the analysis of the classification criteria.

4. Discussion.

Considering all the variables that architectural projects must integrate to become a physical spatial solution, this StyleGAN image generation process is only successful in allowing users to explore new graphic relationships and configurations of recognizable elements in a pixel arrangement. This is generated by the limited relationships between pixels that configure images, given within the limits of what is possible. However, the database frames what the algorithm knows or learns, which means that nothing that has not been previously accepted or selected by the designer or curator can emerge as an outcome, consciously or not.

The use of Machine Learning as part of the creative process affects how the designer communicates and makes design decisions influenced by different aspects of the outcome form. Not only does it affect the images, but it also influences the narrative behind the final results. Creativity is subject to a series of parameters and historical data that is interpreted by technology. Nevertheless, decisions are still under the responsibility of the designer, now acting as the conductor of the orchestra.

The lack of architectural ornaments or historical architectural styles from the chosen images that compose the input database supports the fact that neither of the analyzed synthetic images had a particular or singular element. Generating homogeneous and global criteria of the resulting images that suggest the Chilean architectural house style or characteristics.

Because of the common and homogeneous spatial language of the synthetic images, attributable to the ‘real´ images of the original database, and due the selection of those houses was made by a select group of international curators, along with the website visitors from around the world who recognized relevant architectural design contributions to the field, architectural languages is strongly influenced by their bias, but at the same time, influences the authors of those houses build between 2010 and 2020 selected a ´Project of the Year´ finalists. Building a common and homogeneous spatial and formal architectural language, opening the possibility that a synthetic image could compete in a design or architectural competition with equal footing with ´real’ buildings.

Further research from this investigation could rely on quantifying the relationships between void and solid, advances in the detection of specific architectural elements like columns, and the measurement of the proportion of

the theoretical three-dimensional volumes.

References

"ArchDaily | Broadcasting Architecture Worldwide" : 2020. Available from <www.archdaily.com>.

Baars, B.: 1988, *A Cognitive Theory of Consciousness*, Cambridge University Press.

Bechtel, A.A.a.: 2001, *Connectionism and the Mind*, John Wiley & Sons.

Boden, M.A.: 1998, Creativity and artificial intelligence, *Artificial Intelligence*, **103**(1-2), 347-356.

Brock, A., Donahue, J. and Simonyan, K.: 2018, Large Scale GAN Training for High Fidelity Natural Image Synthesis., *International Conference on Learning Representations.*.

Bátfai, N.: 2011, Conscious Machines and Consciousness Oriented Programming., *arXiv*.

del Campo, M.M.: 2020, A Question of Style - Style, Artificial Intelligence and Architecture, *Architectural Intelligence*.

Karras, T., Laine, S. and Aila, A.: 2019, A Style-Based Generator Architecture for Generative Adversarial Networks, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Leach, N.: 2019, Do Robots Dream of Digital Sleep?, *Proceedings of ACADIA 19:Ubiquity and Autonomy*.

Priore, P., G'omez, A., Pino, R. and Rosillo, R.: 2014, Dynamic scheduling of manufacturing systems using machine learning: An updated review, *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, **28**(1), 83-97.

Shai Shalev-Shwartz, S.B.D.: 2019, *Understanding Machine Learning*, Cambridge University Press.