

***STABLE DIFFUSION* APRENDE DE SEBASTIANO SERLIO: DIBUJO DE ARQUITECTURA CON INTELIGENCIA ARTIFICIAL**

***STABLE DIFFUSION* LEARNS FROM SEBASTIANO SERLIO: ARCHITECTURAL DRAWING WITH ARTIFICIAL INTELLIGENCE**

Elena Merino-Gómez; orcid 0000-0003-4129-4626 UNIVERSIDAD DE VALLADOLID

Fernando Moral Andrés; orcid 0000-0002-5511-8239 UNIVERSIDAD ANTONIO DE NEBRIJA, MADRID

Blanca Querol; orcid 0000-0005-5295-6148 UNIVERSIDAD ANTONIO DE NEBRIJA, MADRID

Pedro Reviriego Vasallo; orcid 0000-0003-2273-1341 UNIVERSIDAD POLITÉCNICA DE MADRID

doi: [10.4995/ega.2024.20332](https://doi.org/10.4995/ega.2024.20332)





El reciente desarrollo de herramientas de Inteligencia Artificial (IA) generativa capaces de generar imágenes a partir de secuencias de texto, está creando oportunidades en muchas disciplinas. La arquitectura no es una excepción y los generadores de texto a imágenes (*text-to-image generators*) se pueden emplear en la representación gráfica. Sin embargo, en la mayoría de los casos no son capaces de generar dibujos arquitectónicos siguiendo el estilo gráfico de autores concretos. Muy recientemente, las herramientas de IA ofrecen al usuario la posibilidad de realizar un reentrenamiento parcial utilizando un pequeño conjunto de imágenes. Esto abre la posibilidad de desarrollar generadores de texto a imágenes con IA personalizados. Sería posible, por ejemplo, recrear la forma de representar arquitecturas de un autor determinado. En este artículo exploramos el potencial de estos generadores personalizados

Durante la última década, la inteligencia artificial (IA) ha logrado avances notables, especialmente en las áreas de *deep learning* y, más recientemente, en la de la IA generativa. Esos avances se han adoptado con gran rapidez y hoy en día están ampliamente difundidas herramientas de inteligencia artificial como ChatGPT o DALL-E, capaces de generar texto e imágenes. La expresión gráfica arquitectónica no será ajena al inmenso impacto esperable en la mayoría de las disciplinas.

Los generadores de texto a imágenes tienen especial relevancia en

utilizando los trabajos de Sebastiano Serlio como estudio de caso.

Los resultados muestran que los generadores de IA personalizados pueden capturar el estilo de Serlio, abriendo el campo de investigación a estudios en profundidad sobre los modos idiosincráticos de la expresión gráfica de arquitecturas a lo largo de la historia.

PALABRAS CLAVE: SEBASTIANO SERLIO, INTELIGENCIA ARTIFICIAL, CUSTOMIZACIÓN, STABLE DIFFUSION

The recent development of generative Artificial Intelligence (AI) tools capable of generating images from text sequences is creating opportunities in many disciplines. Architecture is no exception, and text-to-image generators can be used in graphic representation. However, in most cases, they are not capable of generating architectural drawings following the graphic style of specific authors. Very recently, AI tools offer users the possibility of

arquitectura dada la importancia del dibujo y de los elementos visuales (Ploennings & Berger, 2023). Se han utilizado, por ejemplo, para plantear hipótesis sobre la finalización de arquitecturas inacabadas (Merino-Gómez, et al., 2023) o como herramienta en el proceso creativo de proyectos arquitectónicos (Jaruga-Rodolska, 2022). Sin embargo, el impacto de la IA en el ámbito de la arquitectura ha sido hasta el momento mucho más limitado que en otros sectores. Un ulterior paso, no solo tecnológico, sino también creativo, podría con-

partial retraining using a small set of images. This opens the possibility of developing custom text-to-image generators with AI. For example, it would be possible to recreate the way a particular author represents architectures. In this article, we explore the potential of these custom generators using the works of Sebastiano Serlio as a case study. The results show that custom AI generators can capture Serlio's style, opening the research field to in-depth studies on the idiosyncratic modes of graphic expression in architectures throughout history.

In this article, we explore the potential of these custom generators using the works of Sebastiano Serlio as a case study. The results show that custom AI generators can capture Serlio's style, opening the research field to in-depth studies on the idiosyncratic modes of graphic expression in architecture throughout history.

KEYWORDS: SEBASTIANO SERLIO, ARTIFICIAL INTELLIGENCE, CUSTOMIZATION, STABLE DIFFUSION

During the last decade, Artificial Intelligence (AI) has achieved notable advancements, especially in the areas of deep learning and, more recently, generative AI. These breakthroughs have been rapidly adopted, and today, artificial intelligence tools like ChatGPT or DALL-E, capable of generating text and images, are widely prevalent. These new technologies will have a significant impact across various disciplines. Cultural heritage, and specifically architecture, will not be exceptions.

Text-to-image generators are particularly relevant in architecture given the importance of drawing and visual elements (Ploennings & Berger, 2023). They have been used, for example, to hypothesize the completion of unfinished architectures (Merino-Gómez, et al., 2023) or as a tool in the creative process of architectural projects (Jaruga-



1

Rodolska, 2022). However, the impact of AI in the field of architecture has been much more limited up to this point. A further step, not only technological but also creative, may involve generating unprecedented images created following the graphic expression traits and projection style of any author.

General-purpose AI text-to-image generators are typically trained with vast datasets consisting of text-image pairs, mostly sourced from the internet (Schumann, et al., 2022). This allows them to learn how to generate images for many text sequences, achieving results that are sometimes outstanding. However, when the input text is very specific, general AI tools are unable to produce consistent images. This is because the training set may have no or only a few images paired with that text, and therefore, the generator is unable to produce coherent images. Therefore, generating graphic material with the style of creators whose work has limited presence on the internet proves to be very challenging through AI.

Training a general AI generator with billions of parameters requires a tremendous computational effort, in addition to the availability of a large number of relevant text-image pairs. Therefore, in most cases, it is not feasible for specific application domains. Instead, a pre-trained general model can be fine-tuned using a much smaller set of images and associated text. This allows the development of customized AI generators, opening up new possibilities for their use in any discipline related to graphic arts. In this article, we explore the potential of customized AI generators for architectural drawing by attempting to recreate the style of graphic representations by Sebastiano Serlio.

Text-to-image generators in artificial intelligence and customization

The introduction of DALL-E by OpenAI, with its accessible web interface, has led to an

increase in generating unprecedented images that are created following the graphic expression traits and projection style of any author.

Los generadores de texto a imágenes de IA de uso general se entrenan habitualmente con inmensos conjuntos de datos de pares de texto e imágenes, tomados en la mayoría de los casos de Internet (Schumann, et al., 2022). Esto les permite aprender a generar imágenes para muchas secuencias de texto, logrando resultados que en algunos casos resultan impactantes. Sin embargo, cuando el texto de entrada es muy específico, las herramientas generales de IA no son capaces de producir imágenes consecuentes. Ello se debe a que el conjunto de entrenamiento puede tener ninguna o solo unas pocas imágenes emparejadas con ese texto y, por lo tanto, el generador no es capaz de producir imágenes coherentes. Por tanto, resulta muy difícil la generación mediante AI de material gráfico con el estilo de creadores cuya obra tiene presencia limitada en la red.

El entrenamiento de un generador de IA general con miles de millones de parámetros requiere un enorme esfuerzo computacional, además de la disponibilidad de una gran cantidad de pares de texto e imágenes relevantes. Por lo tanto, en la mayoría de los casos no es factible para dominios de aplicación específicos. En cambio, un modelo general que ha sido previamente entrenado se puede ajustar

utilizando un conjunto mucho más pequeño de imágenes y texto asociado. Esto permite el desarrollo de generadores de IA personalizados, lo que abre nuevas posibilidades para su uso en cualquier disciplina relacionada con las artes gráficas.

En este artículo exploramos el potencial de los generadores de IA personalizados para el dibujo arquitectónico al intentar recrear el estilo de las representaciones gráficas de Sebastiano Serlio.

Generadores de texto-a-imagen en inteligencia artificial y customización

La introducción de DALL-E por parte de OpenAI, permitiendo su acceso a través de una interfaz web fácil de usar, ha llevado a un crecimiento explosivo de los generadores de texto-a-imagen de IA de uso general. El usuario puede introducir un texto y las herramientas generan imágenes que corresponden al texto. Esto se logra mediante el uso de modelos de IA que son capaces de asociar texto con imágenes para luego generar imágenes nuevas pertinentes (Ramesh, et al., 2022).

La popularidad de DALL-E propició el desarrollo de muchos generadores de imágenes de IA, que normalmente se pueden utilizar a través de una interfaz fácil de usar, que no requiere ningún conocimiento de los algoritmos de IA utilizados en los



1. A la izquierda, imagen original del *Libro settimo* (Serlio, 1575, p. 105). Centro y derecha, dos imágenes generadas por Stable Diffusion con la secuencia de texto "Architectural drawing of a building façade, by Sebastiano Serlio"

1. On the left, the original image from the *Libro settimo* (Serlio, 1575, p. 105); center and right, two images generated by Stable Diffusion with the text sequence "Architectural drawing of a building façade, by Sebastiano Serlio"

modelos. Algunos ejemplos de los generadores más ampliamente utilizados son DALL-E, MidJourney, Stable Diffusion y Leonardo AI.

Los modelos de estas herramientas están entrenados para responder a textos de entrada generales y no son capaces de generar imágenes para entradas muy específicas relativas a textos cuyas imágenes correspondientes no están presentes en sus conjuntos de entrenamiento. Para ilustrar este punto, le pedimos a Stable Diffusion que generara dibujos de una fachada al estilo de Serlio. En la figura 1 se observa que el resultado de Stable Diffusion es ajeno a los modos de expresión gráfica de Serlio.

Del mismo modo que Serlio trató de elaborar mediante su material gráfico y textual un análisis de los elementos arquitectónicos de su tiempo, los conjuntos de datos de entrenamiento de inteligencia artificial, basados en textos e imágenes podrían considerarse, de alguna manera, sucesores de aquella metodología (Jaminet, et al., 2022). La naturaleza sistemática del trabajo de Serlio es coherente con la reciente introducción de generadores de imágenes de IA personalizables que se pueden ajustar con un reentrenamiento parcial para reproducir un tipo determinado de imágenes. Estas técnicas ahora forman parte de las opciones de algunos generadores de imágenes de uso general, como Leonardo AI, que permite personalizar Stable Diffusion mediante un pequeño conjunto de datos de imágenes relevantes proporcionadas por el usuario. Posteriormente, el modelo se entrena utilizando las imágenes introducidas para producir un nuevo modelo adaptado al nuevo conjunto de datos. Este proceso es mucho más rápido que en-

trenar un modelo desde cero, algo que no sería factible con un conjunto de datos pequeño.

El dibujo arquitectónico en la época de Sebastiano Serlio

La fortuna de la obra de Sebastiano Serlio se comprende en el contexto de las primeras décadas de la imprenta y de los albores de la asociación de imágenes a los textos que las describen. Los cuadernos de modelos visuales antes de la imprenta eran escasos para la disciplina arquitectónica y, por ello, muy preciados objetos, copiados a mano con tanta fidelidad como era posible (Carpo, 2001).

Así como en el presente se analizan los efectos que los distintos softwares de representación gráfica (Nejadriahi & Arab, 2017) y, recientemente, la IA son capaces de introducir en la creación arquitectónica (Trabucco, 2021), la aparición de la imprenta supuso una revolución inmediata también para la transmisión de la imagen. Las técnicas conjuntas de impresión de texto y gráficos permitieron la introducción de ilustraciones en los tratados de arquitectura. La relativamente sencilla reproducibilidad ofrecía a los autores la oportunidad, no solo de definir gráficamente la teoría, sino la de introducir variantes y presentar combinaciones gráficas de los distintos elementos arquitectónicos. El perfeccionamiento de las técnicas xilográficas y de grabado que podían acompañar a los textos impresos (Bland, 1974) significó la mayor revolución hasta la fecha. En este contexto de innovación técnica empezará a publicarse, en distintos libros, el tratado de arquitectura de Sebastiano Serlio.

explosive growth in general-purpose AI text-to-image generators. Users can input text, and the tools generate images corresponding to the text. This is achieved using AI models capable of associating text with images and then generating relevant new images (Ramesh, et al., 2022). The popularity of DALL-E has spurred the development of many AI image generators, typically accessible through a user-friendly interface that requires no knowledge of the AI algorithms used in the models. Some examples of widely used generators include DALL-E, MidJourney, Stable Diffusion, and Leonardo AI.

These tools' models are trained to respond to general input texts and are unable to generate images for very specific inputs related to texts whose corresponding images are not present in their training sets. To illustrate this point, we asked Stable Diffusion to generate a drawing of a facade in the style of Serlio. In Figure 1, it can be observed that the result from Stable Diffusion does not align with Serlio's modes of graphic expression.

Similar to how Serlio attempted to elaborate on architectural elements in his time through his graphic and textual material, artificial intelligence training datasets based on texts and images could be considered, in a way, successors to that methodology (Jaminet, et al., 2022). The systematic nature of Serlio's work aligns with the recent introduction of customizable AI image generators that can be fine-tuned with partial retraining also known as finetuning to reproduce a specific type of images. These techniques are now part of the options in some general-purpose image generators, such as Leonardo AI, which allows customization of Stable Diffusion with a small set of relevant user-provided image data. Subsequently, the model is trained using the introduced images to produce a new model tailored to the new dataset. This process is much faster than training a model from scratch, which would not be feasible with a small dataset.

Architectural drawing in the time of Sebastiano Serlio

The widespread recognition of Sebastiano Serlio's work is understood in the context of the early decades of the printing press and the emergence of associating images with written texts that describe them. Visual model sketchbooks before the advent of the printing press were scarce in the field of architecture and, therefore, highly prized objects, meticulously copied by hand as faithfully as possible (Carpo, 2001).



Just as the effects of various graphic representation software (Nejadriahi & Arab, 2017), and more recently, AI, on architectural creation are analyzed today (Trabucco, 2021), the advent of the printing press also brought about an immediate revolution in the transmission of images. The combined techniques of printing text and graphics allowed the introduction of illustrations in architectural treatises. The relatively straightforward reproducibility provided authors with the opportunity not only to graphically define theory but also to easily introduce variations and present graphic combinations of various architectural elements. The refinement of woodcut and engraving techniques that could accompany printed texts (Bland, 1974) marked the most significant revolution to date. In this context of technical innovation, Sebastiano Serlio's architectural treatise began to be published in various books.

The 16th century witnessed new purposes in architectural treatises. The loss of illustrations in the transmission of Vitruvius' text and the deliberate absence of images that Alberti intended for his treatise limited its immediate dissemination to highly specialized sectors. However, with the advent of printing and the possibility of accompanying texts with woodcuts and engravings, the deciphering of architectural theory became more accessible, expanding the reception of works to less-educated or non-specialized audiences. Just as present-day AI applications are accessible and user-friendly for the general public, in the works of Serlio and his followers, there was also an expressed intention to be intelligible to a broad audience. In the preamble of his *Quarto libro* (Serlio, 1537) (the first one to be published), Serlio indicates that his work is directed towards readers of both elevated and mediocre understanding ¹.

Another multiplying effect resulting from the use of AI, in addition to its accessibility for non-specialized users, is the introduction of almost infinite variability options, in an augmented version of what happened with treaties and printing from the early 16th century. The architectural treaties of Alberti and Vitruvius, lacking illustrations in their initial editions given to printing, consisted of texts aiming for the canonical definition of architecture. However, their successive versions and the emergence of new treaties led to the relaxation

El siglo XVI fue testigo de nuevos propósitos en la tratadística de Arquitectura. La pérdida de las ilustraciones de Vitruvio en su transmisión textual y la ausencia deliberada de las imágenes que Alberti pretendió para su tratado acotaron su difusión inmediata a sectores altamente especializados. Sin embargo, con la llegada de la imprenta y las posibilidades de acompañar los textos con xilografías y grabados, se incrementaron las posibilidades de desciframiento de la teoría de la arquitectura y se amplió la recepción de las obras a estratos de menor formación o de formación no especializada. Así como en el presente las aplicaciones de IA son accesibles y fáciles de usar para el público en general, en las obras de Serlio y sus epígonos también se expresó en su día el propósito de resultar inteligibles para un público amplio. En el preámbulo de su *Quarto libro* (Serlio, 1537) (el primero que se publicó), Serlio indica que su obra está dirigida a lectores tanto de índole elevada como aquellos de mediocre entendimiento ¹.

Otro efecto multiplicador derivado del uso de la IA, además de su accesibilidad para usuarios no especializados, es la introducción de opciones de variabilidad casi infinitas, en una versión aumentada de lo que ocurrió en relación con los tratados y la imprenta desde principios del siglo XVI. Los tratados de arquitectura de Alberti y Vitruvio, desprovistos de ilustraciones en sus primeras ediciones dadas a la imprenta, consistían en textos cuyo propósito era la definición canónica de arquitectura. Sin embargo, sus sucesivas versiones y la aparición de nuevos tratados significaron la flexibilización de los modelos y la querencia por la diversidad.

La amplia disponibilidad de aplicaciones de IA afectará indudablemente a la producción gráfica de la arquitectura de una manera comparable a los efectos de la imprenta. Las posibilidades que ofrece la IA para generar nuevos modelos visuales, es decir, para producir un número casi infinito de variaciones con poco esfuerzo y coste material, combinadas con la facilidad de transmisión inmediata de los modelos producidos, podrían tener efectos revolucionarios difíciles de predecir.

De toda la obra publicada de Serlio es posiblemente en el *Livre extraordinaire* (Serlio, 1551) y en el póstumo *Il settimo libro d'architettura* (Serlio, 1575) donde más sistemáticamente se aprecian las características de los libros de modelos que otrora se transmitieran de forma artesanal. La vocación de creación mediante la variación se pone de manifiesto en el prólogo del *Livre Extraordinaire*, donde Serlio representa 50 puertas a modo de catálogo con el que "plusieurs en desiroient la copie [de la puerta que construyó para el Cardenal de Ferrara, Hipólito d'Este] pour s'en servir" ². Las variaciones que ofrece, al objeto de "satisfaire à plusieurs appetits d'hommes" ² recurren incluso a "nouvelles fantasies" en un lenguaje que, en algunos modelos, llega a superar el manierismo y alcanza delirantes formas barrocas.

En el inicio de *Il settimo libro d'architettura* Serlio también se refiere a la gran variedad de "habitationi" diseñadas. Serlio distorsiona voluntariamente el canon para la introducción de variantes. Ello se visualiza en las decenas de representaciones publicadas, hasta el punto de convertirse algunos de sus libros, inopinadamente, en una suerte de "pattern



books” cuya repercusión en la obra construida posterior será evidente (Blunt, 1957). Su inmensa y variada audiencia será campo abonado para las versiones, interpretaciones y modificaciones de sus propuestas hasta finales del siglo XVII (Lemerle, 2014), en una senda paragonable a las posibilidades infinitas de alteración que los modelos de IA permiten hoy en día sobre cualquier creación visual.

Metodología

Las obras de Sebastiano Serlio comparten con nuestros propósitos las características de ofrecer un gran abanico de variaciones y una buena legibilidad de las representaciones, de modo que sean susceptibles de alcanzar a una amplia audiencia. Las múltiples propuestas gráficas pueden, por tanto, utilizarse para el entrenamiento. El generador de imágenes utilizado, previamente entrenado, es Stable Diffusion. Se trata de un modelo extensamente difundido y está disponible en abierto, de modo que los resultados de esta investigación podrían ser replicados y ampliados.

Los conjuntos de datos utilizados provienen del *Libro Settimo*, en concreto se seleccionan entre las 24 casas que Serlio denomina “edifici [...] isolati” (Serlio, 1575, p. 1) a la última, denominada por Serlio “casa numero 6”, perteneciente a la serie de “sei palazzi [...] per gran principi e signori”. Se escogen edificios aislados por su mayor facilidad de lectura y se forman dos series para tratar de incluir la mayor heterogeneidad posible. Para realizar el entrenamiento las casas se han agrupado en dos series: una, con 15 fachadas de esquema tripartito, con elementos de torreón en los flancos y, en algunos casos, también en la parte central; la

otra serie, con 14 alzados de composición eminentemente horizontal. La herramienta admite para el entrenamiento un máximo de 40 imágenes.

Para cada uno de estos conjuntos de datos se ha creado un modelo personalizado utilizando el generador de modelos personalizado (fine-tuned models) Leonardo AI con la categoría “edificios” y Stable Diffusion v2.1 como modelo previamente entrenado. La configuración de Leonardo AI y los conjuntos de datos utilizados para el entrenamiento están disponibles públicamente para facilitar la reproducción de nuestros resultados y la experimentación con los modelos 3.

Evaluación

Los modelos personalizados se evaluaron generando un conjunto de imágenes con secuencias de texto (“prompts”) relevantes y comparándolas después con los dibujos originales de Serlio. A continuación, se utiliza un subconjunto de imágenes para analizar los resultados.

En las figuras 2 y 3 se comparan, respectivamente, dos de las quince fachadas originales de Serlio, introducidas para primer entrenamiento de *Il settimo libro* (Fig. 2), con dos imágenes generadas después del re-entrenamiento parcial de Leonardo (Fig. 3). Los edificios elegidos para constituir el conjunto de datos de entrenamiento corresponden a esquemas tripartitos, con torreones laterales y en ocasiones también centrales. Las dos imágenes de la figura 3 se generaron según la siguiente secuencia de texto: “Fachada de un edificio, de Sebastiano Serlio, dos torreones laterales, boceto”; (“a building facade, by Sebastiano Serlio, two side turrets, sketch”). La primera imagen de la

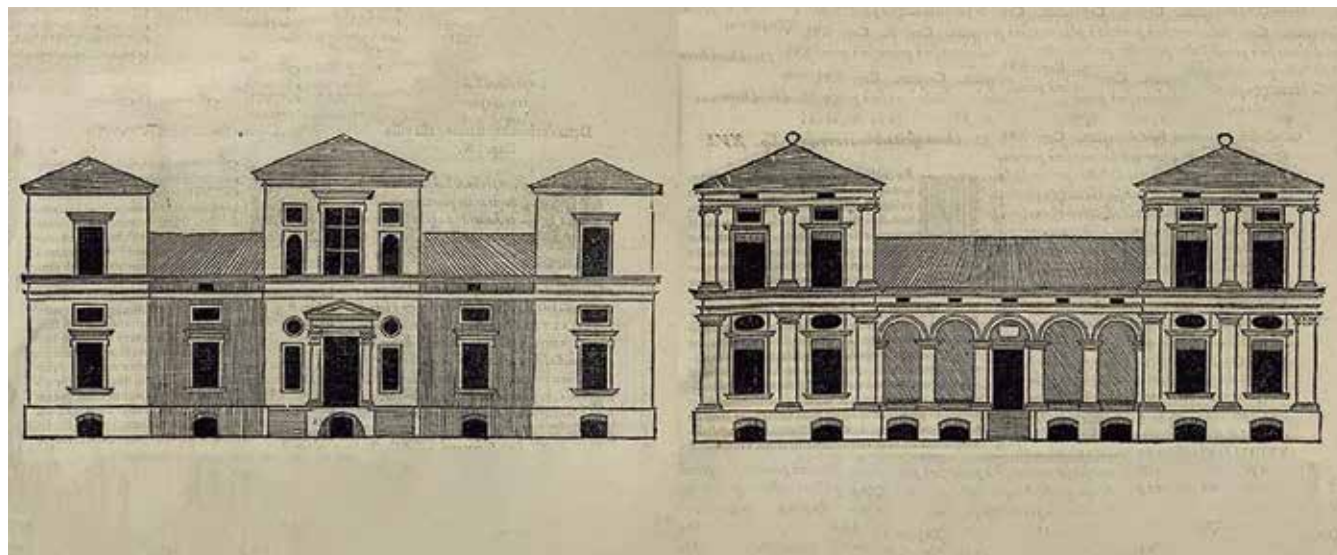
of models and a preference for diversity. Possibly in Serlio’s *Livre extraordinaire* (Serlio, 1551) and in his posthumous *Il settimo libro d’architettura* (Serlio, 1575), the characteristics of the manually transmitted model books are most systematically appreciated from all of Serlio’s published works. The vocation for creating through variation is evident in the prologue of the *Livre Extraordinaire*, where Serlio presents 50 doors in the form of a catalog with which “plusieurs en desiroient la copie [many desired the copy of the door he built for Cardinal of Ferrara, Hippolytus d’Este] pour s’en servir” 2. The variations offered, in order to “satisfy several appetites of men” 2, even resort to “nouvelles fantasies” in a language that, in some models, surpasses Mannerism and reaches delirious Baroque forms.

At the beginning of *Il settimo libro d’architettura*, Serlio also refers to the great variety of “habitationi” designed. Serlio deliberately distorts the canon for the introduction of variants. This is visualized in the dozens of published representations, to the point that some of his books unexpectedly become a kind of “pattern books” whose impact on subsequent built works will be evident (Blunt, 1957). His immense and varied audience will be fertile ground for versions, interpretations, and modifications of his proposals until the late 17th century (Lemerle, 2014), in a path comparable to the infinite alteration possibilities that AI models allow today for any visual creation.

Methodology

The works of Sebastiano Serlio share with our objectives the characteristics of providing a wide range of variations and a high readability of representations, making them capable of reaching a broad audience. The multiple graphic proposals can, therefore, be used for training. The image generator used, previously trained, is Stable Diffusion. It is a widely disseminated model and is available openly, so the results of this research could be replicated and expanded upon.

The datasets used come from the *Settimo Libro*, specifically selected out of the 24 houses that Serlio calls “edifici [...] isolati” (Serlio, 1575, p. 1), up to the last one, referred to by Serlio as “casa numero 6,” belonging to the series of “sei palazzi [...] per gran principi è signori.” Isolated buildings are chosen for their greater readability, and two series are formed to include as much heterogeneity as possible. For training, the



2

houses are grouped into two series: one with 15 facades of a tripartite scheme, featuring turret elements on the flanks and, in some cases, also in the central part; the other series with 14 elevations of predominantly horizontal composition. The tool allows a maximum of 40 images for training.

For each of these datasets, a custom model has been created using the Leonardo AI custom model generator (finetuned models) with the category “buildings” and Stable Diffusion v2.1 as the previously trained model. The configuration of Leonardo AI and the datasets used for training are publicly available to facilitate the reproduction of our results and experimentation with the models 3.

Evaluation

The custom models were evaluated by generating a set of images with relevant text sequences (“prompts”) and comparing them with Serlio’s original drawings. Subsequently, a subset of images is used to analyze the results.

In Figures 2 and 3, two of Serlio’s original facades introduced for the initial training from *Il settimo libro* are compared, respectively (Fig. 2), with two images generated after the partial retraining of Leonardo (Fig. 3). The buildings chosen to constitute the training dataset correspond to tripartite schemes, with lateral and sometimes central turrets. The two images in Figure 3 were generated using the following text sequence: “a building facade, by Sebastiano Serlio, two side turrets, sketch”. The first image was generated using Leonardo AI, while the second was generated using the Alchemy function.

The two images in Figure 3 are generated with complete consistency regarding the training models. The tool has been able to decode the most relevant features of the

figura 3 se generó usando Leonardo AI, mientras que la segunda se generó usando la función Alchemy.

Las dos imágenes de la figura 3 se generan con total consistencia respecto a los modelos de entrenamiento. La herramienta ha sido capaz de decodificar las características más relevantes de las imágenes: elementos arquitectónicos, gama cromática, estilo de dibujo, proporciones y simetría y las ha incorporado a un sinfín de variantes a partir de ellas.

El salto en las posibilidades de generación, donde se hace más patente, de nuevo, es en la comparación de imágenes antes del entrenamiento y las generadas después. Por ejemplo, en la figura 4, se presentan dos imágenes creadas con el *prompt*: “a building facade, by Sebastiano Serlio, two side turrets, sketch”. Se observan dos alzados de índole muy diversa. Solo la simetría y algún elemento arquitectónico considerado de manera independiente se relacionan con Serlio en la primera imagen. En cambio, aspectos fundamentales como la composición arquitectónica, la proporción y las características gráficas distan mucho de los modos de hacer de Serlio. La segunda, generada tras el entrenamiento, es plenamente coherente con los modelos gráficos de Serlio.

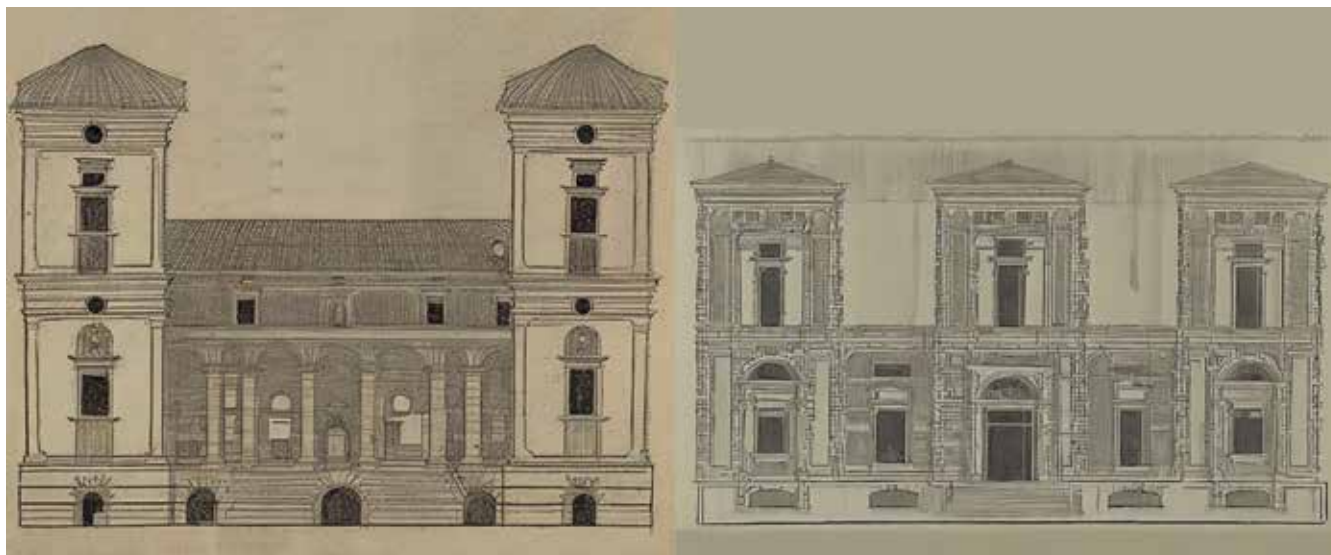
Para verificar la consistencia de la evaluación anterior, se realizó un

ulterior entrenamiento, mediante 14 alzados cuya horizontalidad es la característica principal. Se comparan en las figuras 5 y 6, dos de los alzados originales de Serlio, empleados para el entrenamiento y dos de los resultados obtenidos por la AI, una vez entrenada, e introduciendo el *prompt*: “a one storey facade, by Sebastiano Serlio, sketch”; la primera generada mediante Leonardo AI, y la segunda mediante su herramienta Alchemy.

Se observa en la figura 6 cómo la IA genera alzados de proporciones horizontales, de características muy similares a las imágenes empleadas para el entrenamiento (Fig. 5). A pesar de que el *prompt* exige “a one storey façade” (“una fachada de una planta”), las imágenes generadas siguen el patrón de las empleadas para el entrenamiento, en las que los edificios a veces cuentan con una planta y media, o incluso dos, aunque conservando siempre un predominio de la dimensión horizontal. De ello se puede concluir que prevalece la información gráfica introducida para el entrenamiento sobre la descripción de la secuencia de texto.

Conclusiones

Del análisis de todas las pruebas realizadas podemos concluir que mediante conjuntos de datos muy limitados en número es posible en-



3

trenar Leonardo AI de forma eficaz, de modo que sea capaz de emular las características principales de los gráficos introducidos para el entrenamiento. Los resultados se logran con conjuntos de datos de 15 y 14 ítems respectivamente.

La eficacia del entrenamiento se hace evidente cuando se comparan los resultados antes y después del aprendizaje de la herramienta. Si fuera posible el entrenamiento con la obra completa de Serlio, los resultados serían aún más precisos. Se observa, asimismo, que, en los resultados tras el entrenamiento con imágenes, predominan los datos gráficos sobre las secuencias de texto.

Las posibilidades para la creación arquitectónica que se abren mediante esta metodología son múltiples. Además de las investigaciones en curso sobre la producción de arquitecturas según distintos estilos y autores (Chaillou, 2021), cuya presencia gráfica en la red es muy amplia, es posible extenderlas a autores visualmente menos representados en la red, como se ha demostrado en esta investigación sobre Sebastiano Serlio.

Asimismo, en el campo de la creación arquitectónica, un proyectista podría entrenar la herramienta con sus propios diseños y generar infinitas variantes fácilmente modificando solo las secuencias de texto. ■

Notas

- 1 / “[...] che non pur gli elevati ingegni l’habbiano ad intendere, ma ogni mediocre anchora ne possa esser capace...” (Serlio, 1537, p. v)
- 2 / “...muchos que desean la copia [de la puerta que construyó para el cardenal de Ferrara, Ippolito d’Este] para emplearla” para “...satisfacer los variados gustos de los hombres”. (Serlio, 1551, p. 2)
- 3 / La configuración de la herramienta y los conjuntos de datos para replicar los modelos personalizados están disponibles en https://github.com/revirieg/Serlio_Training_Datasets.

Referencias

- BLAND, D., 1974. *A history of book illustration: the illuminated manuscript and the printed book..* Berkeley: University of California Press.
- BLUNT, A., 1957. *Art and Architecture in France. 1500 to 1700.* Suffolk, Aylesbury: Penguin Books.
- CARPO, M., 2001. How do you imitate a building that you have never seen? Printed Images, Ancient Models, and Handmade Drawings in Renaissance Architectural Theory.. *Zeitschrift für Kunstgeschichte*, 64(2), pp. 223-233.
- CHAILLOU, S., 2021. AI and Architecture. En: I. As & P. Basu, eds. *The Routledge Companion to Artificial Intelligence in Architecture.* London, New York: Routledge: Taylor & Francis Group, pp. 420-441.
- JAMINET, J., ESQUIVEL, G. y BUGNI, S., 2022. *Serlio and Artificial Intelligence: Problematizing the Image-to-Object Workflow.* Singapore, Springer, pp. 3-12.
- JARUGA-RODOLSKA, A., 2022. Artificial intelligence as part of future practices in the architect’s work: MidJourney generative tool as part of a process of creating an architectural form. *Architectus*, 3(71), pp. 95-104.
- LEMERLE, F., 2014. *Architettura, Architecture, textes et images, xviè-xviiè siècles..* [En línea] Available at: http://architecture.cesr.univ-tours.fr/Traite/Notice/ENSBA_Gonse479.asp?param=en [Último acceso: 16 Febrero 2023].

2. Alzados de dos casas originales de *Il settimo libro* de Sebastiano Serlio. A la izquierda, la villa 9 (p. 21); a la derecha, la villa 15 (p. 37)

3. Imágenes generadas con IA tras el entrenamiento. A la izquierda, la imagen generada con Leonardo AI; a la derecha, la imagen generada con la opción Alchemy-Sketch B/W

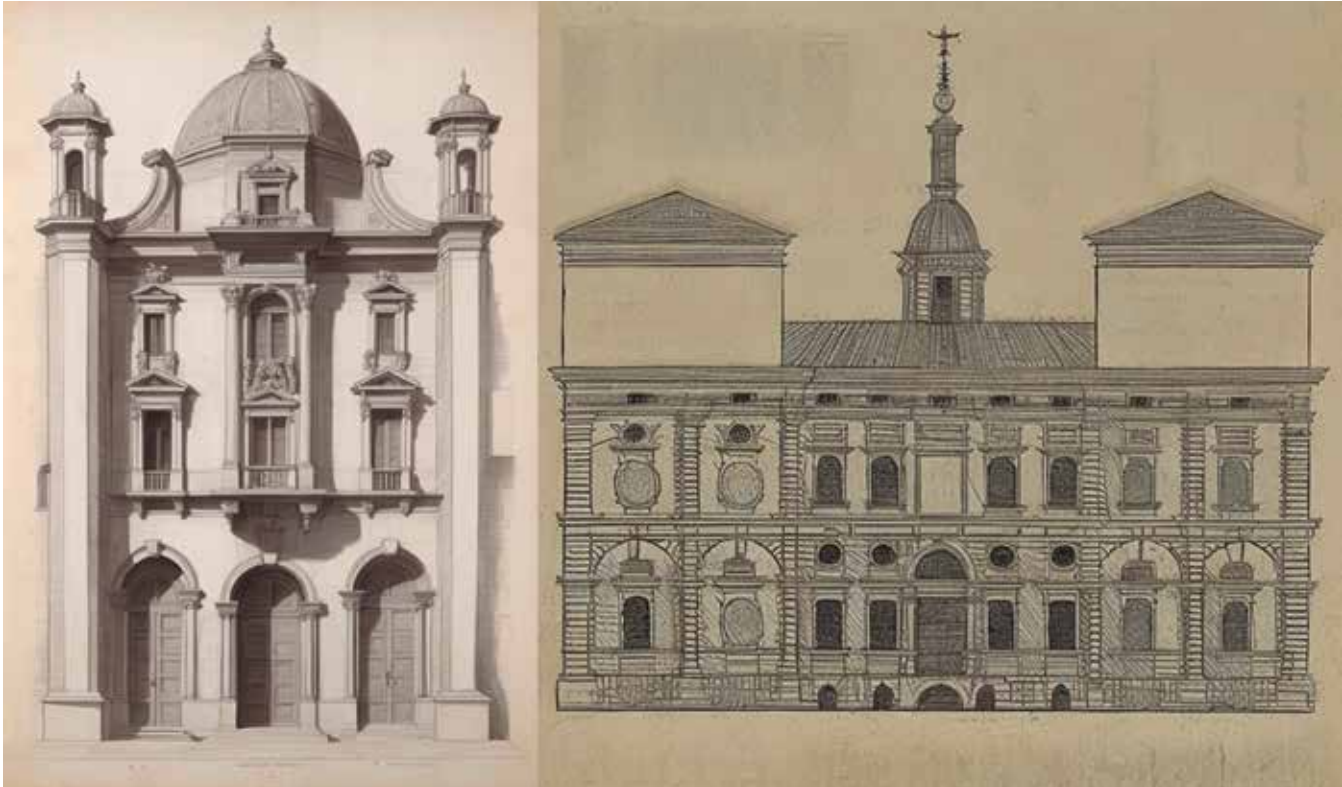
2. Elevations of two unmodified houses from Sebastiano Serlio’s *Il settimo libro*. On the left, villa 9 (p. 21); on the right, villa 15 (p. 37)

3. Two images generated with AI after training. On the left, the image generated with Leonardo AI; on the right, the image generated with the Alchemy-Sketch B/W option

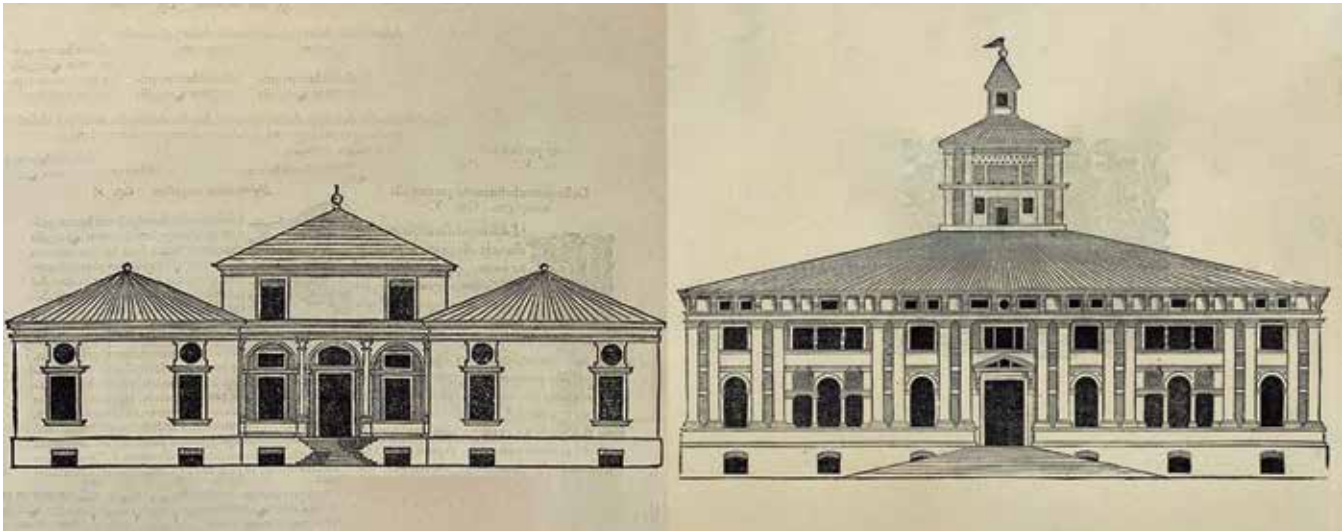
images: architectural elements, color range, drawing style, proportions, symmetry, and has incorporated them into a myriad of variations based on them.

The leap in generation possibilities, where it becomes more apparent again, is in the comparison of images before and after training. For example, in Figure 4, two images created with the prompt “a building facade, by Sebastiano Serlio, two side turrets, sketch” are presented. Two elevations of very diverse nature are observed. Only symmetry and some independently considered architectural elements are related to Serlio in the first image. In contrast, fundamental aspects such as architectural composition, proportion, and graphic characteristics differ significantly from Serlio’s graphical expression. The second image, generated after training, is fully consistent with Serlio’s graphic models.

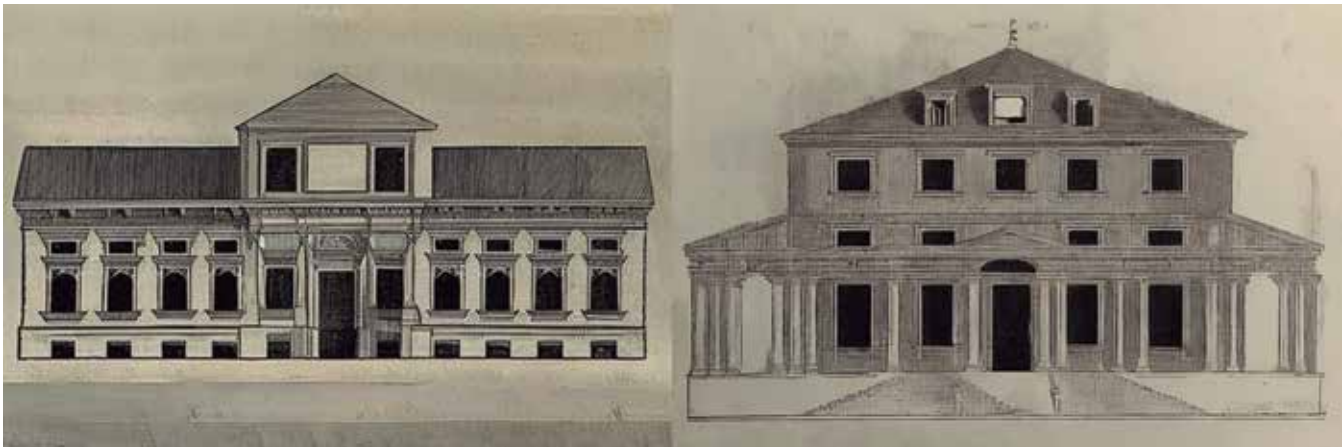
To verify the consistency of the previous evaluation, further training was conducted using 14 elevations characterized by their predominantly horizontal composition. Figures 5 and 6 compare two of Serlio’s original



4



5



6



4. Dos imágenes generadas con IA, ambas con el prompt: “a building facade, by Sebastiano Serlio, two side turrets, sketch”. La de la izquierda se generó antes del entrenamiento; la de la derecha, después

5. Alzados originales de dos casas de *Il settimo libro* de Sebastiano Serlio (1575). A la izquierda, la villa 3 (p. 9); a la derecha, la última casa del volumen, en la que se representa una de las “seis casas para príncipes y señores” (p. 243)

6. Alzados generados mediante AI con el prompt: “a one storey façade, by Sebastiano Serlio, sketch”. A la izquierda, alzado generado mediante Leonardo AI; a la derecha, alzado generado con la opción Alchemy -Sketch B/W

4. Two images generated with AI, both with the prompt “a building facade, by Sebastiano Serlio, two side turrets, sketch.” The one on the left was generated before partial training; the one on the right, after

5. Elevations of two houses from Sebastiano Serlio’s *Il settimo libro*. On the left, villa 3 (p. 9); on the right, the last house in the volume, representing one of the “six houses for princes and lords” (p. 243)

6. Elevations generated using AI with the prompt: “a one storey façade, by Sebastiano Serlio, sketch.” On the left, elevation generated using Leonardo AI; on the right, elevation generated using the Alchemy -Sketch B/W option

- MERINO-GÓMEZ, E., MORAL-ANDRÉS, F. y REVIRIEGO VASALLO, P., 2023. Arquitecturas inconclusas: una perspectiva desde la Inteligencia Artificial. *EGA Expresión Gráfica Arquitectónica*, 28(48), pp. 254-267.
- NEJADRIAH, H. y ARAB, K., 2017. A Study on the Impacts of Computer Aided Design. *International Journal of Architectural and Environmental Engineering*, 11(8), pp. 1054-1058.
- PLOENNINGS, J. y BERGER, M., 2023. AI art in architecture. *AI in Civil Engineering*, 2(8), pp. 1-11.
- RAMESH, A. y otros, 2022. *Hierarchical Text-Conditional Image Generation with CLIP Latents*. s.l.:s.n.
- SCHUMANN, C., BEAUMONT, R., VENCU, R. y GORDON, C., 2022. *LAION-5B: An open large-scale dataset for training*. s.l., s.n., pp. 25278-25294.
- SERLIO, S., 1537. *Regole generali di Architettura...* Venezia: Francesco Marcolini da Forlì.
- SERLIO, S., 1551. *Livre extraordinaire de architecture*. Lyon: Jean de Tournes.
- SERLIO, S., 1575. *Il settimo libro d'architettura, nel quale si tratta di molti accidenti...* Frankfurt am Main: A. Wechel y J. Strada.
- TRABUCCO, D., 2021. Will artificial intelligence kill architects? An insight on the architect job in the AI future. *Journal of Technology for Architecture and Environment*, Volumen 2, pp. 128-132.

Agradecimientos

Este trabajo ha sido posible en parte gracias al proyecto FUN4DATE (PID2022-136684OB-C22) financiado por la Agencia Estatal de Investigación (doi: 10.13039/501100011033).

elevations used for training with two results obtained by AI after training and using the prompt “a one-story facade, by Sebastiano Serlio, sketch.” The first image was generated using Leonardo AI, and the second image was generated using the Alchemy tool.

In Figure 6, it is observed how the AI generates elevations with horizontal proportions, very similar to the images used for training (Fig. 5). Despite the prompt specifying “a one storey façade,” the generated images follow the pattern of those used for training, where the buildings sometimes have one and a half or even two stories, but always maintaining a predominance of the horizontal dimension. From this, it can be concluded that the graphical information introduced during training prevails over the description in the text sequence.

Conclusions

From the analysis of all the conducted tests, we can conclude that it is possible to effectively train Leonardo AI with very limited datasets, enabling it to emulate the main characteristics of the graphics introduced during training. The results were achieved with datasets of only 15 and 14 items, respectively.

The effectiveness of the training becomes evident when comparing the results before and after the tool's learning. If training were possible with Serlio's complete works, the results would be even more precise. It is also observed that, in the results after training with images, graphic data prevails over text sequences.

The possibilities for architectural creation opened up by this methodology are manifold. In addition to ongoing research on the production of architectures according to different styles and authors (Chailou, 2021), whose graphic presence on the internet is extensive, it is possible to extend them to authors who are visually less represented online, as demonstrated in this investigation on Sebastiano Serlio.

Furthermore, in the field of architectural creation, a designer could train the tool with their own designs and generate infinite variations by easily modifying only the text sequences. ■

Notes

- 1 / “[...] not only for elevated minds to understand but also so that even the mediocre can grasp it...” (Serlio, 1537, p. v)
- 2 / “...many who desire a copy [of the door he built for the Cardinal of Ferrara, Ippolito d'Este] to use it” to “...satisfy

the varied tastes of men.” (Serlio, 1551, p. 2)

3 / The configuration of the tool and the datasets to replicate the custom models are available at https://github.com/revirieg/Serlio_Training_Datasets

References

- BLAND, D., 1974. *A history of book illustration: the illuminated manuscript and the printed book*. Berkeley: University of California Press.
- BLUNT, A., 1957. *Art and Architecture in France. 1500 to 1700*. Suffolk, Aylesbury: Penguin Books.
- CARPO, M., 2001. How do you imitate a building that you have never seen? Printed Images, Ancient Models, and Handmade Drawings in Renaissance Architectural Theory. *Zeitschrift für Kunstgeschichte*, 64(2), pp. 223-233.
- CHAILLOU, S., 2021. AI and Architecture. En: I. As & P. Basu, eds. *The Routledge Companion to Artificial Intelligence in Architecture*. London, New York: Routledge: Taylor & Francis Group, pp. 420-441.
- JAMINET, J., ESQUIVEL, G. & BUGNI, S., 2022. *Serlio and Artificial Intelligence: Problematizing the Image-to-Object Workflow*. Singapore, Springer, pp. 3-12.
- JARUGA-RODOLSKA, A., 2022. Artificial intelligence as part of future practices in the architect's work: MidJourney generative tool as part of a process of creating an architectural form. *Architectus*, 3(71), pp. 95-104.
- LEMERLE, F., 2014. *Architettura, Architecture, textes et images, XVIè-XVIIè siècles*. [En línea] Available at: http://architectura.cesr.univ-tours.fr/Traite/Notice/ENSBA_Gonse479.asp?param=en [Último acceso: 16 Febrero 2023].
- MERINO-GÓMEZ, E., MORAL-ANDRÉS, F. & REVIRIEGO VASALLO, P., 2023. Arquitecturas inconclusas: una perspectiva desde la Inteligencia Artificial. *EGA Expresión Gráfica Arquitectónica*, 28(48), pp. 254-267.
- NEJADRIAH, H. & ARAB, K., 2017. A Study on the Impacts of Computer Aided Design. *International Journal of Architectural and Environmental Engineering*, 11(8), pp. 1054-1058.
- PLOENNINGS, J. & BERGER, M., 2023. AI art in architecture. *AI in Civil Engineering*, 2(8), pp. 1-11.
- RAMESH, A. y otros, 2022. *Hierarchical Text-Conditional Image Generation with CLIP Latents*. s.l.:s.n.
- SCHUMANN, C., Beaumont, R., Vencu, R. & Gordon, C., 2022. *LAION-5B: An open large-scale dataset for training*. s.l., s.n., pp. 25278-25294.
- SERLIO, S., 1537. *Regole generali di Architettura...* Venezia: Francesco Marcolini da Forlì.
- SERLIO, S., 1551. *Livre extraordinaire de architecture*. Lyon: Jean de Tournes.
- SERLIO, S., 1575. *Il settimo libro d'architettura, nel quale si tratta di molti accidenti...* Frankfurt am Main: A. Wechel & J. Strada.
- TRABUCCO, D., 2021. Will artificial intelligence kill architects? An insight on the architect job in the AI future. *Journal of Technology for Architecture and Environment*, Volumen 2, pp. 128-132.

Acknowledgments

This work has been possible in part thanks to the FUN4DATE project (PID2022-136684OB-C22) funded by the Spanish Agencia Estatal de Investigación (doi: 10.13039/501100011033).