

Methodological Implementation of StyleGans Algorithms and Its Change of Paradigm in the Education, Practice and Role of Designers

Tomás Vivanco^{1,2}, Antonia Valencia¹, Philip Yuan²

¹ Pontifical Catholic University of Chile

tvivanco@uc.cl

arvalencia@uc.cl

² Tongji University, China

philipyuan007@tongji.edu.cn

Abstract. The integration of Artificial Intelligence algorithms into computational design processes promote a human-machine collaboration, transforming the role of conventional designers into meta-designer as a product of this collaboration. Specifically, StyleGAN's algorithms offer a novel approach to experimenting with shapes from previously designed images of products or objects. This article presents the application of a methodology for image experimentation to designers without previous knowledge of computation and programming. Each of the steps was developed through different cases for different speculative object production, using artificial intelligence algorithms, and reflecting - in an applied way - the designer's role as curator and co-creator of the creative process in conjunction with computing.

Keywords: Design AI, StyleGAN, Computer Graphics, Design Education, Computational Design.

1 Introducción

Los recientes avances en Ciencias de la Computación permiten la integración de nuevas herramientas procedimentales en los procesos de diseño asistido por computador. Construyendo una relación simbiótica entre diseñadores y computación (Scafà, Marconi, Germani, 2020), fomentando la exploración de formas emergentes, construyendo un espacio creativo multidimensional para navegar dentro de un espacio latente.

Conceptualmente, las ideas existen gracias a que podemos representarlas (Suwa, Tversky, 2002), a su vez, la representación implica el uso de herramientas o métodos conocidos o entrenados que hacen explícita la idea para comunicarla. Una nueva idea, que podría ser entendida como creativa,

que no se puede entender ni describir mediante operaciones lógicas. Materializándose, como novedad, cuando esta puede sorprender a un observador o usuario. Por tanto, la capacidad de reconocimiento del atributo de novedad recae en el destinatario de la idea.

Desde la perspectiva del creador, las ideas son producidas por conexiones que no son necesariamente racionales o lógicas originadas principalmente por tres fuentes de información (Glăveanu, Tanggaard, 2014). El primero es el conocimiento previo o genético heredado, el segundo son los sentimientos y percepciones, o sentidos, y finalmente, el conocimiento adquirido durante su vida.

Estas tres fuentes de información son nuestra base de datos masiva, el lugar desde donde exploramos todo el conocimiento o información que alimenta el proceso de diseño creativo único. Desde esta perspectiva, nos preguntamos,

¿es posible explorar los límites de nuestra creatividad a través de la generación de bases de datos de información o imágenes utilizando herramientas computacionales?

Si la información que alimenta la creatividad pudiera ser procesada y organizada para alimentar un proceso de diseño procedimental. En ese caso, almacena las posibilidades de generar nuevas formas y diseños a través de imágenes o referencias formales previamente conocidas y con las que operamos consciente e inconscientemente (Shalev-Shwartz, 2014). Entonces, ¿puede un algoritmo navegar u operar dentro de la creatividad?

Este artículo expone a través de una serie de experimentos de diseño computacional, la exploración de una relación colaborativa y creativa entre diseñadores y algoritmos de Machine Learning, con resultados formales aplicados enmarcados en un ámbito reflexivo sobre el origen de las ideas y espacios imaginarios desde donde diseñamos.

1.1 Redes Generativas Antagónicas (GAN).

Generative Adversarial Networks (GANs) es una técnica de Inteligencia Artificial generativa, basada en un algoritmo en red que reconoce y genera relaciones entre una gran cantidad de información de datos, de manera similar a como funciona el cerebro (Vivanco, Valencia, Yuan, 2021). Permitiendo generar imágenes sintetizadas (StyleGAN) o artificiales, con una resolución muy similar a las imágenes "reales". La información de entrada se procesa en capas de neuronas, reconociendo patrones y generando asociaciones entre ellos, para finalmente generar un resultado. Si el resultado es incorrecto o similar al original, la red aprende de él para generar una nueva imagen, asegurando que su resultado sea nuevo en comparación con la base de datos original.

Las GAN están compuestas por dos redes neuronales que 'compiten' para producir estas imágenes originales. Hay un generador ascendente que genera imágenes y un discriminador descendente que evalúa esas imágenes. De alguna manera, el generador es el artista y el discriminador es el crítico o curador (Leach, 2019), donde el primero siempre tiene que sorprender al segundo. En este proceso de competencia, el generador debe producir imágenes realistas hasta que el discriminador no pueda distinguirlos del conjunto de datos reales.

1.2 Generador GAN

El Generador y Discriminador son dos de los componentes básicos que permiten el funcionamiento de GAN. Ambos componentes son modelos de Redes Neuronales que están en constante refuerzo:

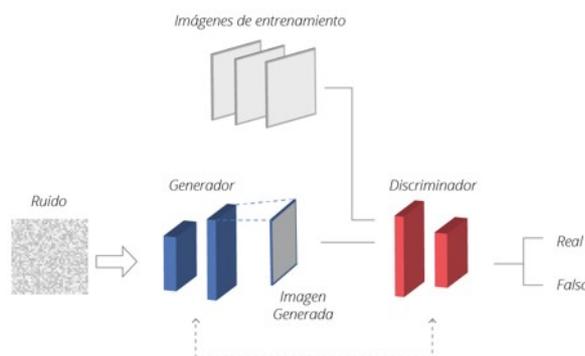


Figura 1. Diagrama del flujo del Generador y discriminador de una Red Generativa Antagónica. Fuente: Elaboración de los autores.

La función del Generador (Figura 1) es producir resultados falsos a partir de ruido aleatorio o de una entrada aleatoria, que como punto de partida son datos reales. El discriminador es el algoritmo que distingue los datos reales de los falsos antes de volver a ingresar esos datos en el generador. Cada flujo completo que realiza el algoritmo dentro de su arquitectura es definida como una época o *epoch*. Con las cuales se puede regular y controlar el proceso de aprendizaje del algoritmo a partir de la base de datos original.

1.3 StyleGAN

Presentado por Nvidia en 2018, ocupa la estructura de una Red Antagonista Generativa (GAN) con el propósito principal de generar imágenes de rostros (Smith, 2020), a partir de una base de datos de imágenes de rostros reales modificadas con ruido e interferencia. Dos redes neuronales interactúan

constantemente, donde una propone nuevas opciones y la otra evalúa su nivel de autenticidad. Similar al proceso en el que los humanos aprenden un estilo formal específico o elementos geométricos (del Campo, Manning, 2020), las redes neuronales se entrenan con tantas imágenes como sea posible (como mínimo mil) para reconocer las diferencias entre elementos y estilos.

Supongamos que el proceso de computar una gran cantidad de imágenes para reconocer motivos (Priore, 2001) y producir nuevas imágenes elegantes con esos patrones es similar al proceso humano de observar, clasificar y crear. En ese caso, las computadoras a través de algoritmos de inteligencia artificial podrían volverse, de alguna manera, tan creativas como los humanos o al menos convertirse en una herramienta de apoyo para navegar a través de la creatividad de una manera extendida y colaborativa.

El aprendizaje humano y el aprendizaje automático aplicado al análisis de imágenes son procesos similares basados en la detección de patrones significativos. Con la diferencia de que el aprendizaje basado en IA es un algoritmo automatizado (Shalev-Shwartz, 2014) sin saber cuál es la finalidad simbólica, cultural y funcional del objeto. Por lo tanto, no puede establecer relaciones o asociaciones funcionales con otros objetos ni generar resultados fuera de los datos entrenados de entrada.

1.4 Meta-diseñadores

Los procesos de diseño convencionales basados en objetos se enfocan en la expresión formal de ideas, que pueden cambiar dependiendo de la incidencia de las fuerzas y reglas detrás de cada proceso. Definir -o diseñar- esas reglas y las relaciones entre las fuerzas, supone el cambio de un proceso de diseño orientado a objetos hacia un proceso de diseño orientado a sistemas (Vivanco, 2016), donde los diseñadores pueden alimentar un sistema con reglas definidas para generar y navegar a través de una gran cantidad de expresiones formales emergentes.

Existe un cambio de escala que cuestiona las prácticas de diseño actuales basadas en diseñar mono-soluciones, elevando la idea de un diseño orientado hacia uno orientado al sistema por medio de un proceso generativo para producir resultados emergentes esperados e inesperados, donde los diseñadores se convierten en curadores en un proceso de meta-diseño.

Las herramientas y procesos de diseño generativo pueden convertirse en aliados clave de los diseñadores para abordar los desafíos computacionales y creativos emergentes, al establecer los entornos de diseño donde las formas no son más que el resultado de las relaciones de diferentes elementos y fuerzas dentro de ese ecosistema (Johnson, 2001). El diseño de códigos con inteligencia artificial implica un territorio exploratorio de resultados predichos e

imprevistos; implica ajustar hiperparámetros y objetivos (Hutter, Lücke, Schmidt-Thieme, 2015) para que, finalmente, habilite un proceso de diseño en una conversación hombre-máquina.

2. Metodología y objetivos.

Siguiendo una metodología experimental, se describe el proceso de generación de imágenes gráficas utilizando algoritmos de inteligencia artificial de diferentes objetos. El cual se estructura en cinco pasos (Figura 2). secuenciales y recursivos; cada uno de ellos explicado a través de estudios de casos aplicados desarrollados en una clase de pregrado de estudiantes de diseño, sin ningún conocimiento previo en programación e Inteligencia Artificial.

En primer lugar, la recopilación de un conjunto de datos de imágenes reales de diferentes objetos (guitarras, lepidópteros, animales, pipas, robots y partituras). En segundo lugar, un proceso experimental con el algoritmo de GAN para probar el conjunto de datos. En tercer lugar, diferentes entrenamientos de StyleGAN ayudan a ajustar el conjunto de datos y la configuración del algoritmo. Cuarto, la definición de la selección y criterios de las imágenes de resultado. Quinto y último la postproducción de los resultados finales.

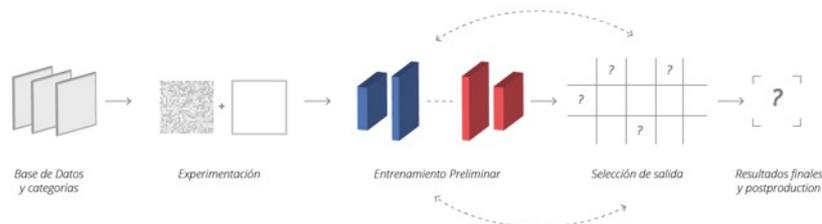


Figura 2. Cinco pasos de la metodología de investigación descrita. Fuente: Elaboración de los autores.

Cada base de datos fue un estudio de caso específico que sometido a todos los pasos metodológicos hasta llegar al resultado final. Sin embargo, con el fin de describir la capacidad de variación y desarrollo de diferentes alternativas como resultados de esta metodología este artículo presenta en cada paso una base de datos de objetos diferente.

2.1 Proceso.

Paso 1. Generación de la Base de datos.

Es fundamental elaborar una base de datos organizada, limpia y robusta y luego pre procesarla, asegurando que tengan un fondo estandarizado, color, eliminación de elementos no deseados y representativos.

Una buena clasificación (Figura 3) de base de datos contribuye a definir elementos funcionales, categorías y subcategorías para unificarla y hacerla más efectiva (Deverall, 2017). Esto puede entenderse como el primer proceso de curaduría, de donde surgirán todas las imágenes sintéticas.



Figura 3. Clasificación de instrumentos de cuerda: Imágenes de vista frontal de los instrumentos y sus componentes morfológicos. Fuente: Elaboración de los autores.

Estudio de caso: Lepidoptera

En busca de una especie poco probable de Lepidopteras se recolectó una base de datos de 137 polillas y 369 mariposas. Además, para tener una mayor diversidad cromática y de texturas (Dubey, 2020), una base de datos compuesta por 317 siluetas de polillas creadas a partir de imágenes de texturas (Figura 4) con el modelo textures_dtd.

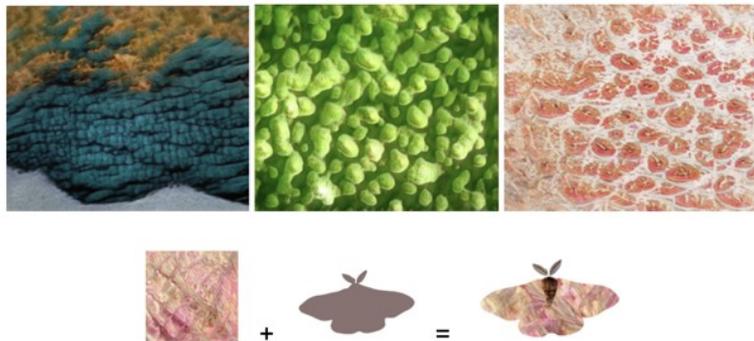


Figura 4. Texturas generadas con el modelo textures_dtd.

Paso 2. Experimentación.

Este paso desarrolla la experimentación con un modelo StyleGAN previamente entrenado para generar rostros humanos, utilizando la base de datos de código abierto para experimentar y lograr resultados con la base de datos recopilada. Con la intención de encontrar y utilizar modelos que se ajusten a las

proyecciones y exploración futuras, buscando explorar nuevas imágenes de salida.

Estudio de caso: instrumentos de cuerda

Múltiples iteraciones en el diseño de instrumentos de cuerda han llevado a la perfección de su forma y materialidad, generando una versión estandarizada y aparentemente terminada como proceso diseño.



Figura 5. Experimentación inicial de nuevas imágenes de instrumentos de cuerda.
Fuente: Elaboración de los autores.

Al generar una línea de productos conceptuales el proceso elimina el sesgo morfológico de los instrumentos de cuerda estándar (Figura 5). Al brindar una base de datos diversa y extensa, los resultados coinciden en aquellas partes básicas que son comunes a todos los instrumentos, pero generan nuevas expresiones formales. De esta forma, la curaduría de los resultados da derecho a seleccionar aquellos ejemplos que destaquen por su novedad formal. Sin embargo, no necesariamente por su potencial para ser un instrumento de alta calidad, sino por su capacidad para desafiar los diseños estándar existentes.

Paso 3. Entrenamiento preliminar:

A partir de la base de datos pre-entrenada, se modifican los hiperparámetros de épocas, muestra utilizada para entrenar y muestra para evaluar los resultados. La resolución o calidad del resultado entrenado dependerá en gran medida de la calidad de la base de datos y del número de pasos de entrenamiento.

La navegación dentro del espacio latente ofrece infinitas posibilidades de imágenes sintéticas, las cantidades de posibilidades implica la curaduría en la evaluación y selección de los resultados finales. El truncamiento y la distancia de muestreo (Figura 6), marcan la variabilidad de las imágenes de salida dentro de los vectores todas sus dimensiones. Un incremento en la distancia de muestreo aumenta la diversidad, distinguiendo significativamente una imagen de otra dentro en el espacio latente. El truncamiento, por otro lado, determina qué tan "realista" será la imagen de salida. Cuanto más cerca esté el número de 0, más se aproxima al conjunto de datos original.

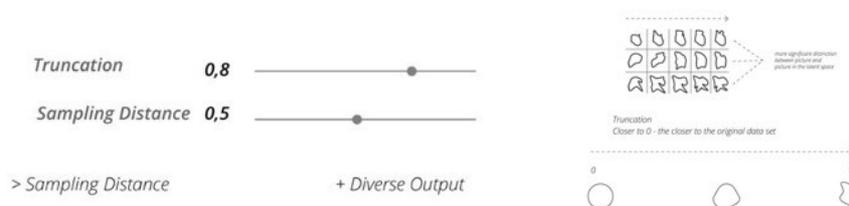


Figura 6. Diagrama de transformación de resultados cambiando los parámetros de distancia de muestreo y truncamiento. Fuente: Elaboración de los autores.

Estudio de caso: Animales Mecha

Consiste en un modelo para producir colecciones de modelos híbridos de robots y mamíferos. El cruce generativo de estilos depende de una base de datos de animales compuesta por 1179 imágenes, y otra de robots de 867 imágenes. Durante el proceso de desarrollo de los modelos, se configuraron los siguientes parámetros:

Mamíferos: 1179 imágenes, 2350 pasos, truncamiento 0,8 Distancia de muestreo 0,5.

Robots: 867 imágenes, 2169 pasos, Truncamiento 0.8 Distancia de muestreo 0.5.

Resultados del modelo:

Aunque las imágenes sintéticas regeneradas eran muy similares a los datos de entrada originales, el algoritmo recordaba el modelo de entrada original. Al disminuir la influencia de la base de datos sobre el modelo pre-entrenado para generar rostros humanos (Figura 7), quizás la diversidad de animales no fue suficiente para influir en la salida de deformación de rostros y animales.

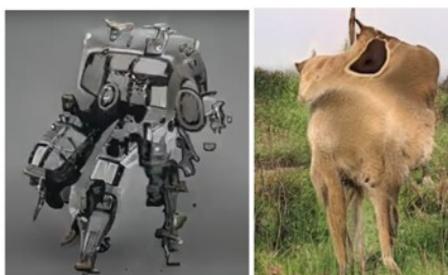


Figura 7. Resultados iniciales de una imagen de salida Generative Adversarial Network StyleGAN de robots entrenados usando un modelo de automóvil y mamíferos usando un modelo facial. Fuente: Estudiantes OPR Diseño por Simulación Escuela de Diseño UC.

Esta limitación generada por el sesgo del modelo debido a la baja diversidad de la base de datos, y el número de pasos, no fue suficiente para generar cambios significativos. Observar la evolución en el modelo previamente entrenado refleja cómo la base de datos influye en el resultado mediante su curación.

Paso 4. Selección de salida:

La visión de la colaboración humano-máquina por medio del control de hiperparámetros permite regular la pérdida de aprendizaje y evitar la memorización de la base de datos durante el entrenamiento por parte del algoritmo, fortaleciendo el buen desempeño durante la producción de imágenes (Sbai, Elhoseiny, Bordes, LeCun, Couprie, 2018).

Estudio de caso: Pipas

El objetivo es entrenar un modelo de IA y explorar formas emergentes para de pipas, considerando sus partes fundamentales. Para ser funcional, debe tener tener al menos una boquilla, al menos un quemador y al menos un conducto. Par ello, un modelo de detección de objetos fue entrenado para posteriormente analizar las imágenes artificiales, seleccionando las imágenes que cumplen con los criterios previamente establecidos.

Las imágenes sintéticas pasaron el filtro de detección de objetos como si fueran objetos reales. Esto debido a los elementos funcionales de las imágenes "reales" (Figura 8), también incorporados en las imágenes artificiales, engañando al algoritmo de detección de objetos, previamente entrenado con imágenes reales, que ahora reconoce lo artificial como real.



Figura 8. Imágenes generadas que pasan los criterios de detección de objetos.
Fuente: Estudiantes OPR Diseño por Simulación Escuela de Diseño UC.

Paso 5. Resultados finales y postproducción:

El uso de Machine Learning como parte del proceso creativo afecta la forma en que el diseñador se comunica y toma decisiones de diseño influenciadas por diferentes aspectos de la forma de resultado. Como hemos visto hasta ahora, afecta a las imágenes, pero también influye en la narrativa detrás de los resultados finales.

Estudio de caso: peluches.

Después de aplicar los pasos anteriores de la metodología, el objetivo es generar nuevas familias y relatos de peluches. Permitiendo la generación virtual de elementos "reales", tales como fondos y planetas (Figura 9), así como las descripciones obtenidas con las narrativas escritas de Inferkit (InferKit, 2020). Los resultados abrieron de nuevos imaginarios, más allá del peluche como producto, sugiriendo escenarios especulativos (del Campo, Manning, 2020).



Figura 9. Mundos generados con Spade Coco y figuras generadas con StyleGan.
Fuente: Estudiantes OPR Diseño por Simulación Escuela de Diseño UC.

Estudio de caso: Partituras generativas

Entendiendo las partituras como expresión gráfica musical. Una especie de código sinestésico capaz de transcribir y viajar a través de los sentidos y sensaciones, abriendo la posibilidad de experimentar y encontrar la vía sensorial óptima para la creación en el campo de la inteligencia artificial (Brown, 2009). Bajo esta premisa nació la hipótesis del experimento; generar melodías con inteligencia artificial a través de imágenes.

Para ello, se desarrolló una base de datos variada en estilos musicales, siguiendo reglas claras y estrictas con el fin de obtener resultados sencillos y fáciles de reproducir de forma audible.

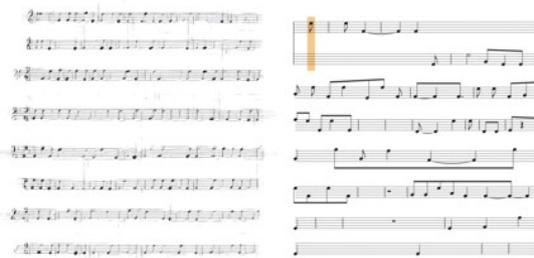
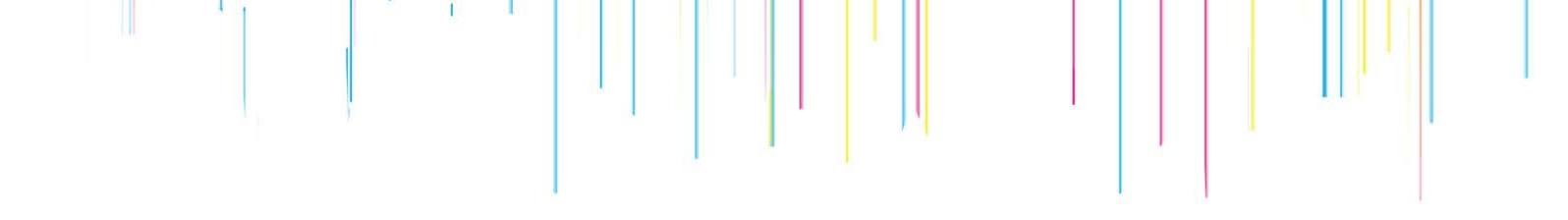


Figura 10. Postproducción de imágenes generadas. A la derecha, una partitura sintética reproducible. Fuente: Estudiantes OPR Diseño por Simulación Escuela de Diseño UC.

Las imágenes resultantes (Figura 10) fueron analizadas con un lector de música, interpretando las partituras creadas, reconociendo correctamente la



secuencia de tiempo definida de los compases y la posición relativa de las notas en ellos (Collins, 2008). Si bien las partituras generadas no son del todo legibles y coherentes, se genera un elemento que se vuelve interpretable por reproductores de música virtuales.

Discusión

La distinción entre computación y computarización (Terzidis, 2006) no se trata siempre en disciplinas creativas. Ciertamente, todavía existe una distancia con la computación ya que su potencial no es reconocido como una extensión creativa de nuestras capacidades humanas. En cierto modo, las herramientas del lápiz y el papel permiten experimentar y explorar diferentes posibilidades formales sin ninguna barrera de entrada. Los avances en el desarrollo de varias plataformas computacionales para disminuir la complejidad de la barrera de entrada (Dorning, 2020) para utilizar herramientas computacionales, democratizan el acceso a usuarios sin ninguna experiencia en codificación.

El uso de Machine Learning como parte del proceso creativo afecta la forma en que el diseñador se comunica y toma decisiones de diseño influenciadas por diferentes aspectos de la forma del resultado. No solo afecta las imágenes, sino que también influye en la narrativa humana detrás de los resultados finales. La creatividad está sujeta a una serie de parámetros y datos históricos que son interpretados por la tecnología. Además, las imágenes sintéticas basadas en IA se producen solo con una base de datos predefinida, no puede hacer nada fuera de lo definido en esa base de datos. Incluso si se definen las fronteras más amplias de la considerada "inspiración" en un proceso creativo.

Agradecimientos

Estudiantes del curso Diseño por Simulación de la Escuela de Diseño de la Pontificia Universidad Católica de Chile. Cristóbal Valenzuela y Runway ML. Grupo Doctoral Digital Futures de Tongji University.

Referencias

- Brown, A. R., & Sorensen, A. (2009). Interacting with generative music through live coding. *Contemporary Music Review*, 28(1), 17-29.
- Collins, N. (2008). The analysis of generative music programs. *Organised sound*, 13(3), 237-248.
- del Campo, M., Manninger, S. (2020). A Question of Style - Style, Artificial Intelligence and Architecture. *Architectural Intelligence*.
- el Campo, M., Manninger, S. (2020). Towards Hallucinating Machines - Designing with Computational Vision. *International Journal of Architectural Computing*.

- Deverall, J., Lee, J., Ayala, M. (2017). Using Generative Adversarial Networks to Design Shoes: The Preliminary Steps. <http://cs231n.stanford.edu/reports/2017/pdfs/119.pdf>
- Dorning, J., 2020. My First Generative ML MODEL With Runway ML—Aimon. Medium. <https://uxdesign.cc/my-first-generative-ml-model-with-runway-ml-aimon-9dccb5fca50d>
- Dubey, A.(2020). AI Assisted Apparel Design Book. <https://arxiv.org/pdf/2007.04950.pdf>
- Glăveanu, V. P., & Tanggaard, L. (2014). Creativity, identity, and representation: Towards a socio-cultural theory of creative identity. *New Ideas in Psychology*, 34, 12-21.
- Hutter, F., Lücke, J. & Schmidt-Thieme, L. (2015). Beyond Manual Tuning of Hyperparameters. *Künstl Intell* 29, 329–337.
- InferKit. (2020). Text Generation - InferKit. <https://inferkit.com/docs/generation>
- Johnson, S. (2001). *Emergence: The connected lives of ants, brains, cities, and software*. New York: Scribner.
- Leach, N. (2019). Do Robots Dream of Digital Sheep?. *ACADIA 19:Ubiquity And Autonomy Conference Proceedings*.
- Priore, P., Fuente, D., Gomez, A. Puente, J. (2001). Dynamic Scheduling of Manufacturing Systems with Machine Learning.. *Int. J. Found. Computation*.
- Sbai, O., Elhoseiny, M., Bordes, A., LeCun, Y., & Couprie, C. (2018). Design: Design inspiration from generative networks. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*.
- Scafà, M. Marconi, M. Germani, M. (2020) A critical review of symbiosis approaches in the context of Industry 4.0, *Journal of Computational Design and Engineering*, Volume 7, Issue 3. Pages 269–278.
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge: Cambridge University Press.
- Smith. R. (2020) Artificial Intelligence Makes Blurry Faces Look More Than 60 Times Sharper
- Suwa, M., & Tversky, B. (2002, April). External representations contribute to the dynamic construction of ideas. In *International Conference on Theory and Application of Diagrams* (pp. 341-343). Springer.
- Terzidis, K. (2006). *Algorithmic architecture*. Amsterdam; Boston: Architectural Press.
- Vivanco, T, Valencia, A., Yuan, P. (2021) Spatial Findings On Chilean Architecture Stylegan Ai Graphics. *Proceedings of the 26th CAADRIA Conference - Volume . pp. 251-260*
- Vivanco, T. (2016). *Design and Fiction as Instruments for Social Development*. Distributed Design Book. Ediciones UC.