

Diseño arquitectónico usando redes convolucionales

Architectural design using convolutional networks

Héctor Julián Cardona Montoya

Facultad de ingenierías, Universidad Tecnológica de Pereira, Pereira, Colombia

Correo-e: julian.cardona@utp.edu.co

Karem Giovanna Parra Moreno

Facultad de ingenierías, Universidad Tecnológica de Pereira, Pereira, Colombia

Correo-e: karem.parra@utp.edu.co

Alberto Andrés Ospina López

Facultad de ingenierías, Universidad Tecnológica de Pereira, Pereira, Colombia

Correo-e: alberto.ospina@utp.edu.co

Resumen— En el siguiente trabajo se presenta un software que permite al usuario crear una imagen de tipo arquitectónica y estructural completamente nueva a partir de la disposición y ubicación de diferentes figuras de colores y formas básicas. Este software fue desarrollado con redes neuronales convolucionales que básicamente son una variación de un perceptrón multicapa y que son muy efectivas para tareas de visión artificial, como clasificación y segmentación de imágenes, además, se usó una red discriminadora que determina si las muestras que son ingresadas como imágenes son reales comparándolas con los mismo ejemplos del entrenamiento.

Palabras clave— Dominio, GAN (redes generativas antagónicas), imagen, red discriminadora, redes neuronales convolucionales.

Abstract— The following work presents software that allows the user to create a completely new architectural and structural image from the arrangement and location of different figures of colors and basic shapes. This software was developed with convolutional neural networks that basically are a variation of a multilayer perception and that are very effective for tasks of artificial vision, as classification and segmentation of images, in addition, a discriminating network was used that determines if the samples that are entered as images are real comparing them with the same examples of the training.

Key Word — convolutional neural networks, discriminatory network, domain, GAN (antagonistic generative networks), image.

I. INTRODUCCIÓN

Al ver las figuras 1 y 2 es fácil para los espectadores reconocer que ambas son similares en cuanto a la disposición y distribución de los objetos dentro de las pinturas, sin embargo, es evidente que ambas tienen diferencias con respecto al color y ambiente con el que se realizó la obra. En primer lugar, tenemos un paisaje de invierno realizado por Claude Monet llamado “Haystacks snow effect” (Figura 1), esta pertenece a una serie de pinturas denominada “Almiares” donde se puede visualizar que el artista capturó, desde diferentes perspectivas y momentos del día, los mismos montones de trigo formados en el campo después de la cosecha, otra de las pinturas de esta serie se observa en la Figura 2 llamada “Haystacks en of summer”.

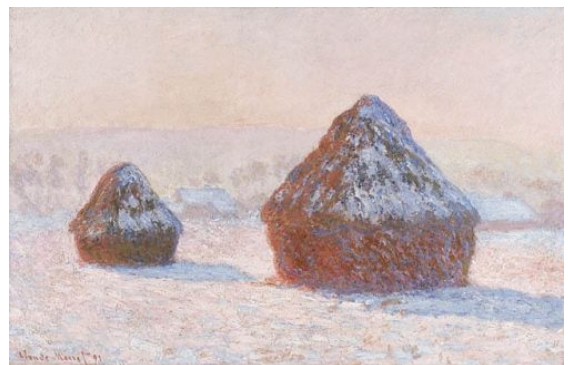


Figura 1. Pajares con efecto de nieve de Claude Monet



Figura 2. Pajares al final del verano de Claude Monet

Es posible lograr imaginar un cambio entre fotografías similar al que se planteó con anterioridad. Podemos razonar sobre las diferencias de color y estilísticas entre estas dos imágenes, y por lo tanto imaginarnos cómo podría ser una escena si la "tradujéramos" de un conjunto al otro. En este trabajo, presentamos un método que puede aprender a hacer lo mismo: capturar las características esenciales de una figura que es dibujada con figuras y colores básicos, con la disposición de los objetos de la imagen, realizar una traducción con imágenes nuevas para dar origen a una nueva imagen a partir de un entrenamiento previo realizado por el algoritmo.

II. CONTENIDO

1. Funcionamiento básico del algoritmo.

Traducir una imagen a otra imagen es una tarea en la cual se crea una versión sintética de la imagen original, pero aplicando una modificación o cambio específico de manera controlada, algunos ejemplos comunes de esto es cambiar imágenes de caballos por cebras o cambiar paisajes de verano por paisajes de invierno.

Para lograr esto existen los modelos supervisados que requieren de un gran conjunto de datos de ejemplos emparejados con los cuales poder entrenar el modelo, sin embargo, conseguir este conjunto de datos no es nada sencillo y puede ser una tarea muy costosa e incluso imposible debido a la necesidad de lograr encontrar imágenes prácticamente iguales en las que los únicos cambios son los que queremos aplicar.

Por suerte existen los modelos no supervisados los cuales implican técnicas de capacitación automática sin la necesidad de entrenar con imágenes de ejemplo emparejadas a través de la arquitectura GAN, una de estas técnicas es CycleGAN, gracias a esta los modelos entrenan de manera no supervisada utilizando un conjunto de imágenes del dominio origen y destino que no es necesario que estén relacionadas.

Como dijimos CycleGAN hace uso de la arquitectura GAN la cual se compone de un modelo generador y otro discriminador los cuales hacen parte de una especie de juego o competición en la cual el modelo generador toma una imagen del dominio origen como entrada y genera nuevas imágenes para el dominio destino, al mismo tiempo el modelo discriminador toma una imagen del dominio destino y predice o clasifica si dicha imagen es real o es generada (falsa). Al hacer uso de esta arquitectura GAN ambos modelos se entrenan mutuamente con el fin de que el modelo generador pueda crear mejores imágenes que engañe al discriminador y este último se entrena con el fin de reconocer y clasificar mejor las imágenes generadas. Además, CycleGAN implementa la capacitación de dos modelos generadores y dos discriminadores con el fin de que los procesos de traducción de imágenes sean en los dos sentidos, es decir, un generador toma imágenes del dominio origen como entrada y produce imágenes para el dominio destino, y el otro generador toma imágenes del dominio destino como entrada y genera imágenes para el dominio origen, y a su vez los modelos discriminatorios se encargan de detectar las imágenes generadas en ambos dominios.

Esto por si solo sería suficiente para crear imágenes aceptables en cada dominio, sin embargo no es suficiente para generar traducciones de las imágenes de entrada ya que al recibir una imagen de entrada y procesarla a través del primer método generador para luego procesarla por el segundo método generador (el proceso inverso) se debería obtener la imagen inicial, pero para que esto suceda es necesario añadir una extensión a la arquitectura llamada "consistencia del ciclo" la cual añade una pérdida adicional para medir la diferencia entre la salida del segundo generador y la imagen original (y también al inverso), esto con el fin de actuar como guía a los métodos generadores hacia la traducción de imágenes.

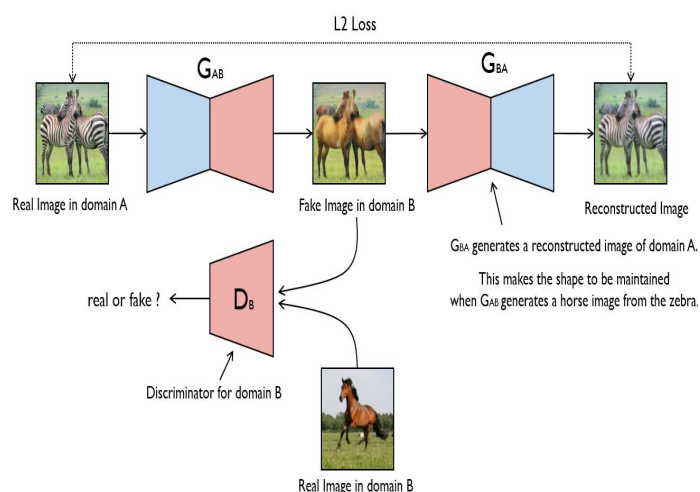


Figura 5. Abstracción de CycleGAN

3. Desarrollo del algoritmo.

Básicamente el algoritmo se basa en la unión de una red neuronal convolucional la cual se encarga de la aplicación de filtros a la imagen que se le envía a través del canvas del HTML.

La red convolucional envía un resultado el cual es enviado a una red neuronal discriminadora que dice si se parece o no a un edificio.

En caso de que no, la red convolucional aplicará nuevos filtros y hará los cambios de los pesos respectivos.

En caso contrario la red neuronal aplicará nuevos filtro para tratar de obtener un resultado mejor optimizado, en caso de que el resultado sea peor. La red terminará su procesos y guardará los pesos.

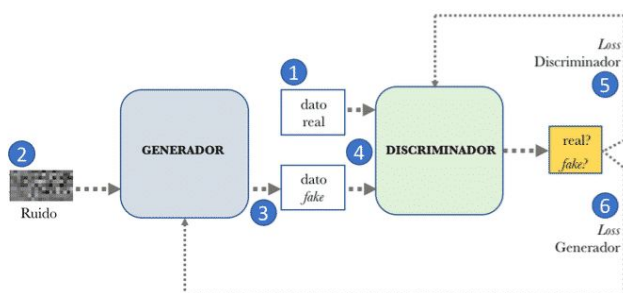


Figura 4.Arquitectura

III. CONCLUSIONES

Del anterior desarrollo se concluye:

- Las redes convolucionales son efectivas para el tratamiento de imágenes, más comúnmente y funcional para la visión artificial, como clasificación y segmentación de imágenes.
- Las redes GAN en especial son una manera de atribuirle cierta capacidad de imaginación a las máquinas y aunque aún no son perfectas pueden lograr resultados bastante buenos tanto así que se podrían utilizar en muchas áreas científicas y de ingeniería, por ejemplo GANs que predicen cómo se comportaría partículas, o GANs que generan registros falsos de pacientes tan buenos como los reales para analizar el funcionamiento de un medicamento.

REFERENCIAS

- [1] (2019). Wikipedia: Redes neuronales convolucionales. Recuperado de

https://es.wikipedia.org/wiki/Redes_neuronales_convolucionales

- [2] Hao, K. (2019). MIT Technology Review: Bienvenidos a la hermosa y terrorífica era de las GAN. Recuperado de: <https://www.technologyreview.es/s/10789/bienvenidos-la-hermosa-y-terrorifica-era-de-las-gan>

- [3] (2019) Wikipedia: Almiarés. Recuperado de: <https://es.wikipedia.org/wiki/Almiarés>

- [4] Jun-Yan, Z. Park, T. Isola, P. Efros, A. A.(2018). Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. Recuperado de: <https://arxiv.org/pdf/1703.10593.pdf>

- [5] Brownlee, J. (2019). Machine Learning Mastery: A Gentle Introduction to CycleGAN for Image Translation. Recuperado de: <https://machinelearningmastery.com/what-is-cyclegan/>

- [6] Gabriela Gonzalez.(2019). Qué son y cómo funcionan las GAN, esas redes neuronales capaces de crear rostros de personas que no existen. Recuperado de: <https://www.genbeta.com/a-fondo/que-como-funcionan-gan-esas-redes-neuronales-capaces-crear-rostros-personas-que-no-existen>

