

Generative Adversarial Networks

Omar Sanseviero

osanseviero@gmail.com

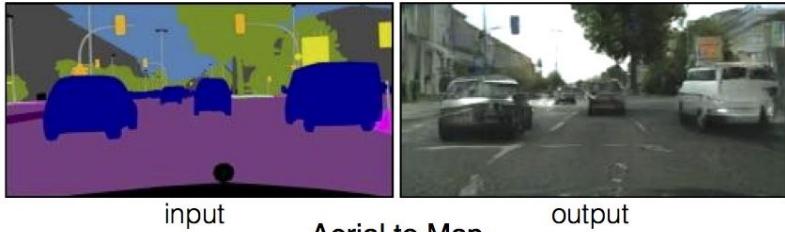
*You might not think that programmers are artists, but
programming is an extremely creative profession. It's
logic-based creativity.* - John Romero



pix2pix (Berkeley, 2017)

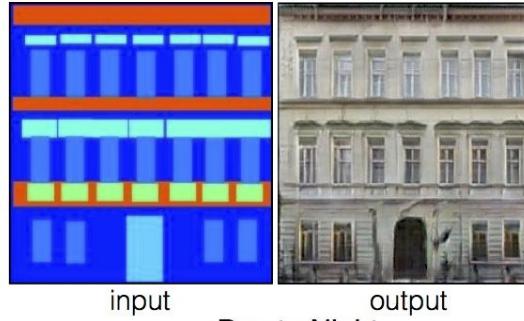
- Traduce imagen a imagen en el mismo dominio

Labels to Street Scene



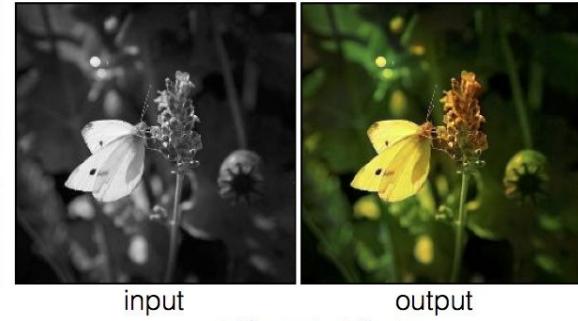
input

Labels to Facade



input

BW to Color



input

Aerial to Map



input

Day to Night



input

Edges to Photo

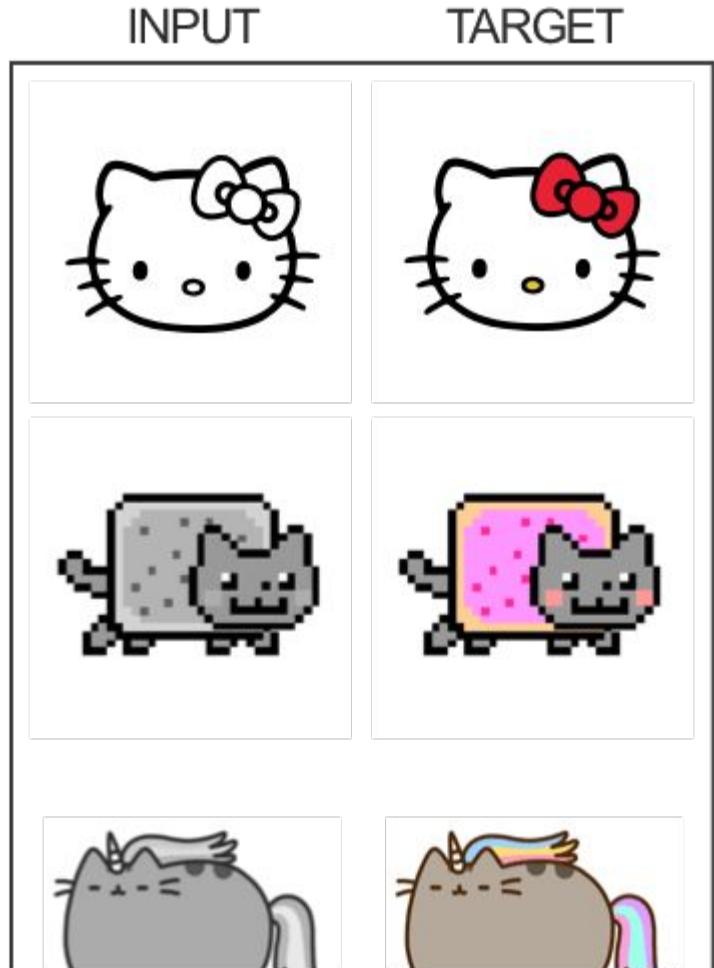


input

output

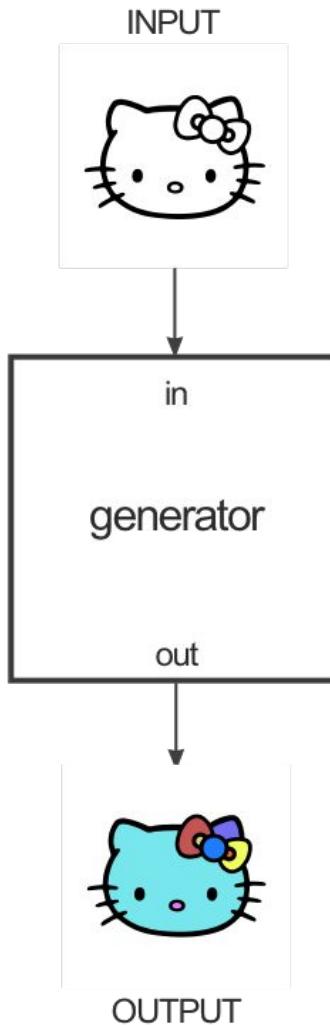
pix2pix (Berkeley, 2017)

- Usa GAN condicional (cGAN)
- Tiene un **discriminador** y un **generador**
- El generador convierte la entrada en la salida
- El discriminador intenta distinguir entre el dataset real y el generado



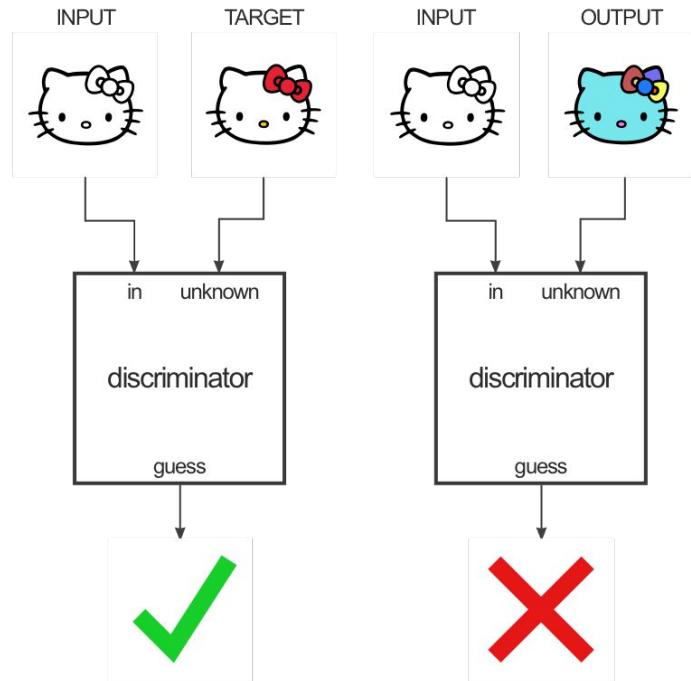
pix2pix (Berkeley, 2017)

- Usa GAN condicional (cGAN)
- Tiene un **discriminador** y un **generador**
- El generador convierte la entrada en la salida
- El discriminador intenta distinguir entre el dataset real y el generado



pix2pix (Berkeley, 2017)

- Usa GAN condicional (cGAN)
- Tiene un **discriminador** y un **generador**
- El generador convierte la entrada en la salida
- El discriminador intenta distinguir entre el dataset real y el generado



pix2pix (Berkeley, 2017)

Background removal



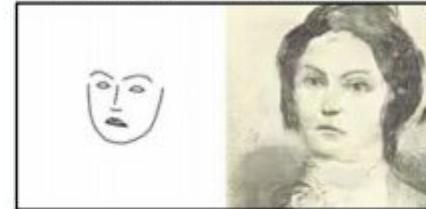
by Kaihu Chen

Palette generation



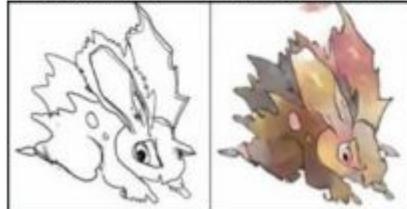
by Jack Qiao

Sketch → Portrait



by Mario Klingemann

Sketch → Pokemon



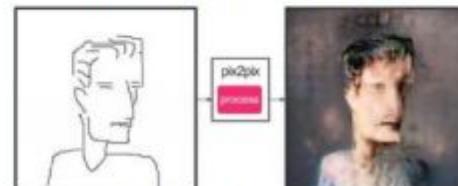
by Bertrand Gondouin

“Do as I do”



by Brannon Dorsey

#fotogenerator

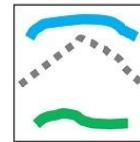
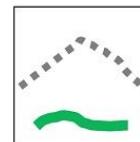
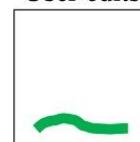


sketch by Yann LeCun

iGAN (Berkeley, 2016)



User edits



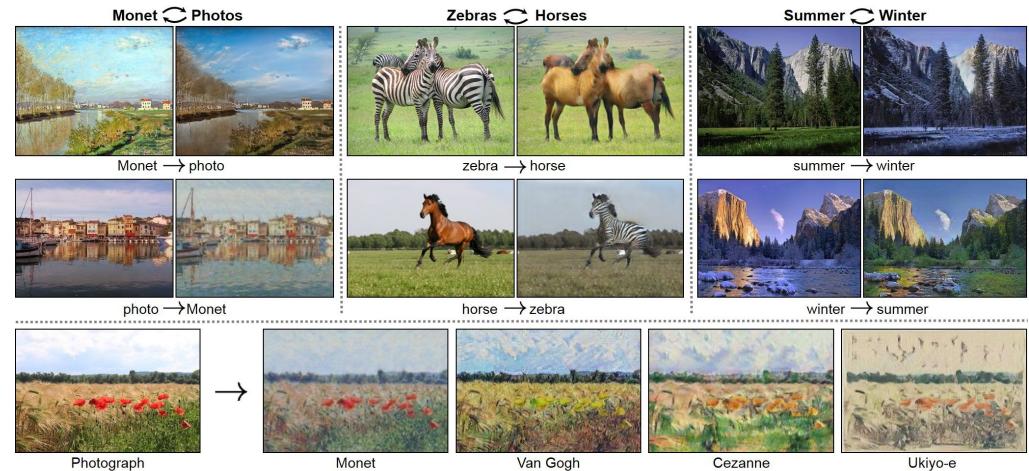
Generated images



— Color

— Sketch

CycleGAN - imagen a imagen



CycleGAN - Cambio de temporada

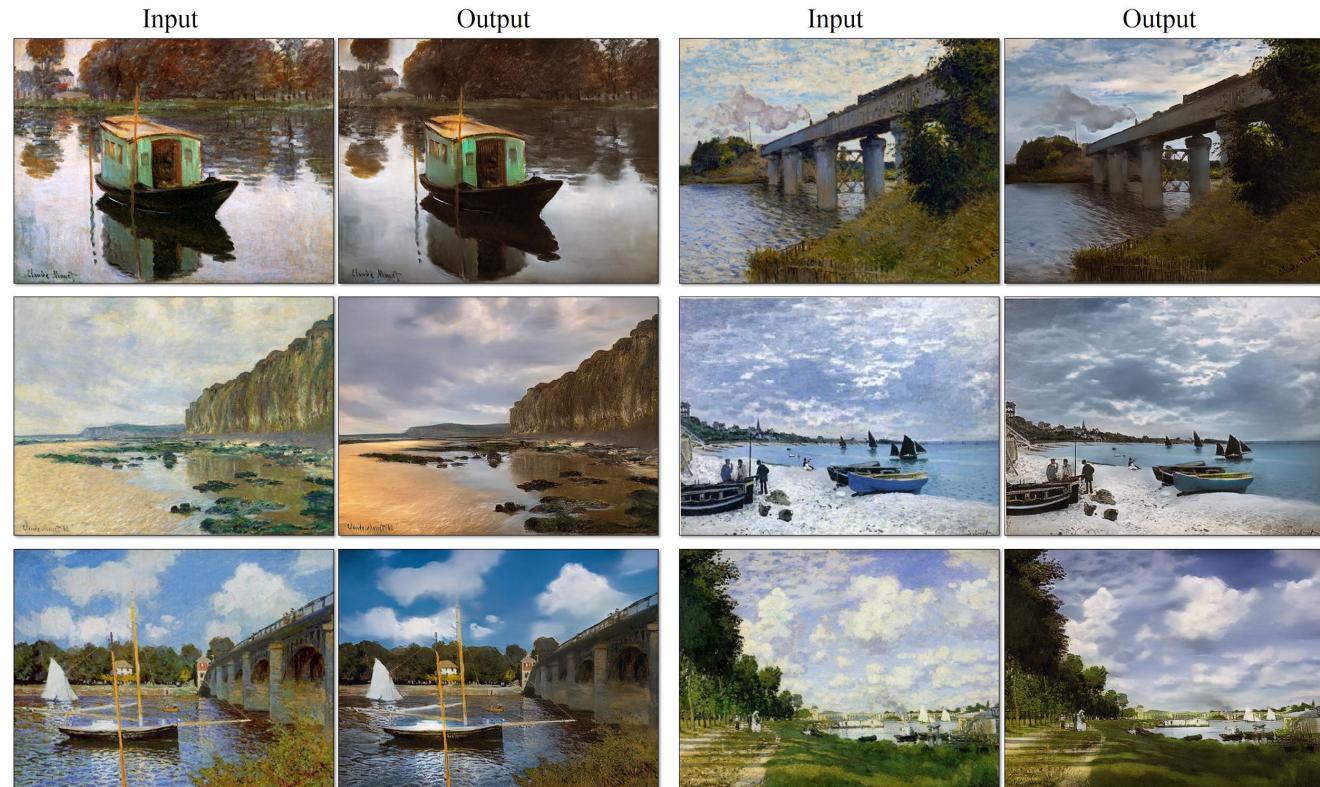


winter Yosemite → summer Yosemite



summer Yosemite → winter Yosemite

CycleGAN - Pintura a foto



CycleGAN - No es perfecto



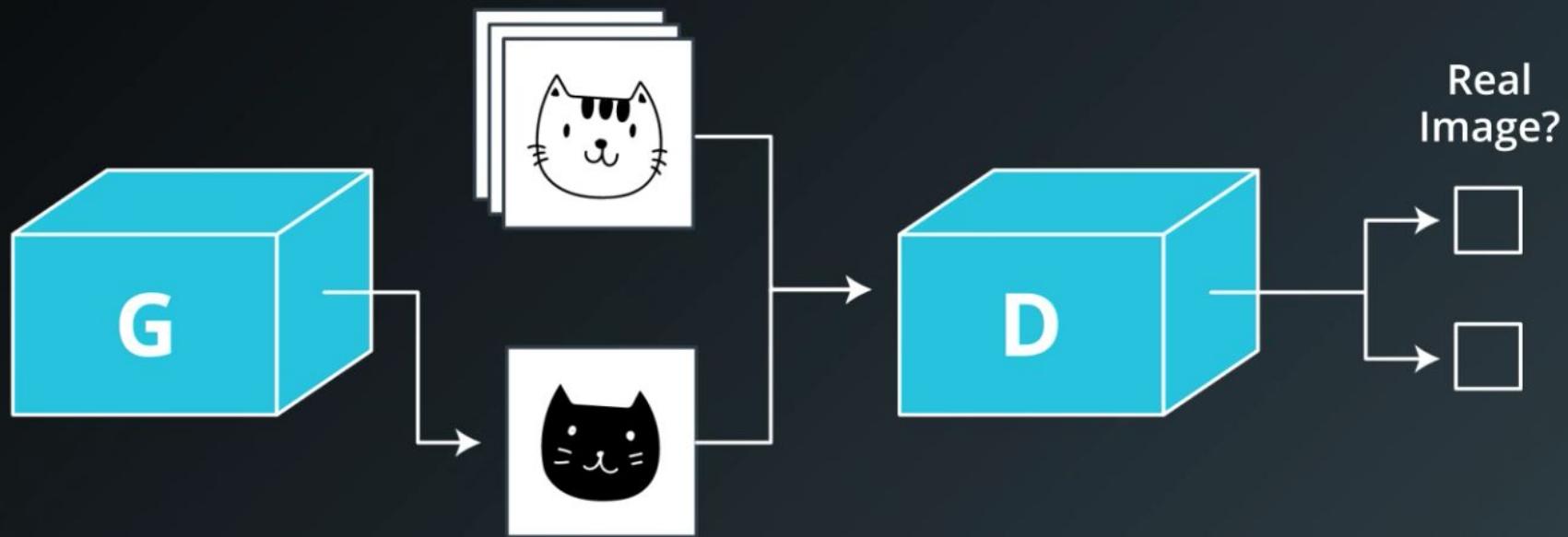
Aplicaciones

- Aprendizaje por imitación
- Simulaciones de experimentos químicos
 - Mucho más económico en vez de simulaciones con monte carlo
- Generar resultados adversariales que permiten entrenar mejor
 - Los GANs pueden intentar engañar a modelo de ML durante el entrenamiento

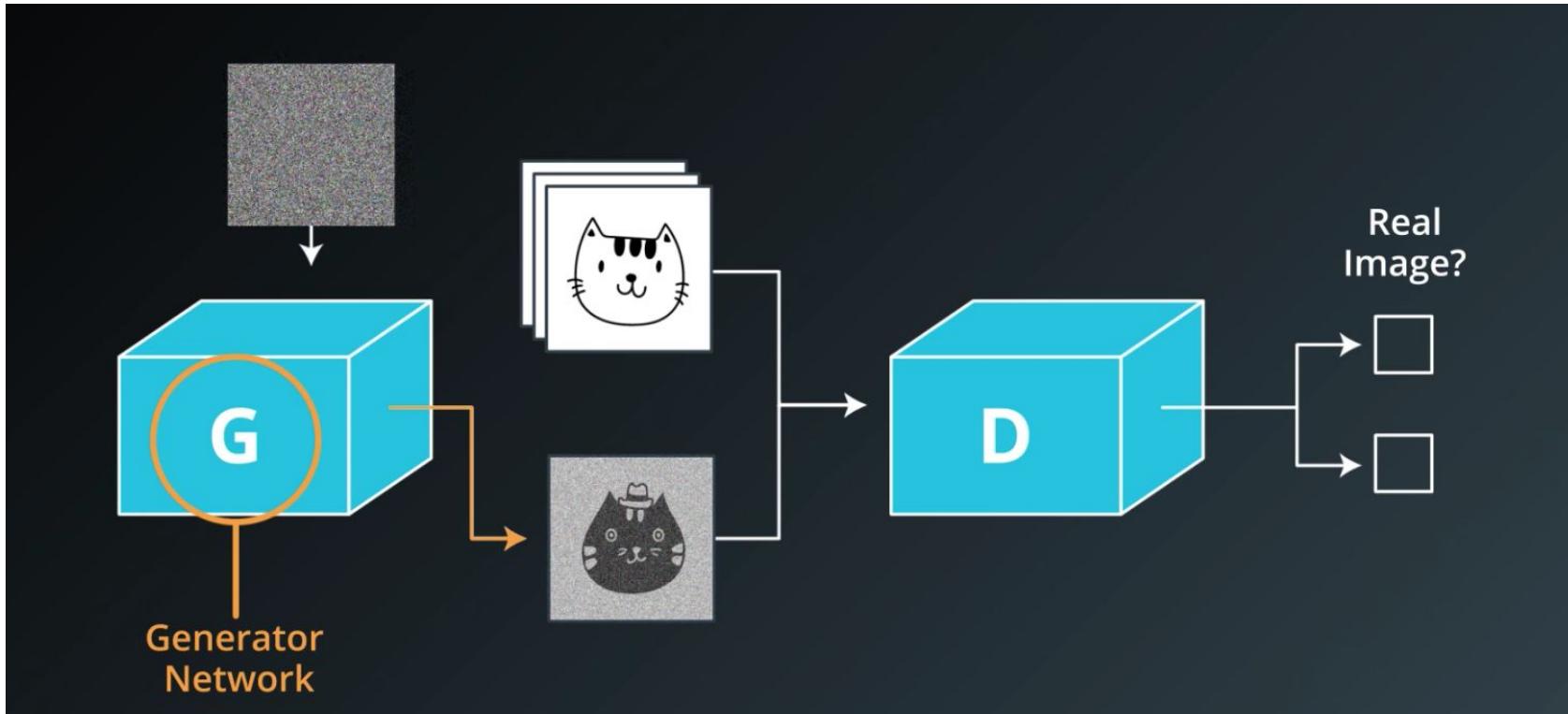
¿Cómo funciona? (Université de Montréal, 2014)

- Con RNNs podemos generar - generaban una palabra a la vez
- Podemos hacer lo mismo con imágenes - generamos un pixel a la vez
- ¿Cómo generamos una oración o una imagen completa de una vez?

¿Cómo funciona?

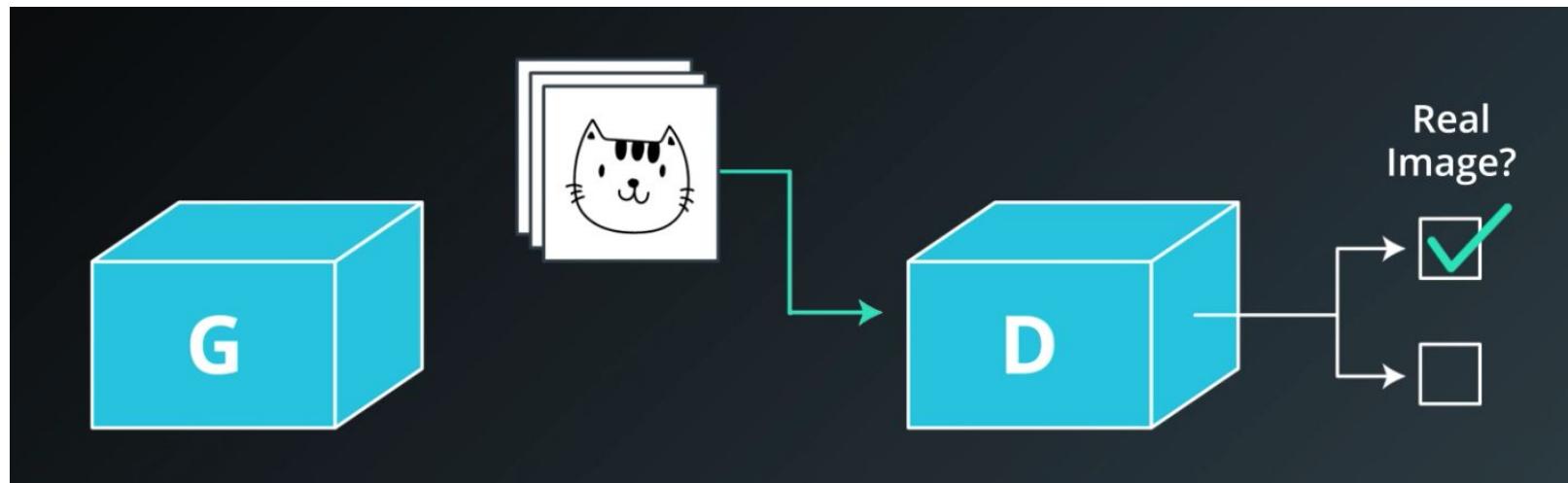


¿Cómo funciona?

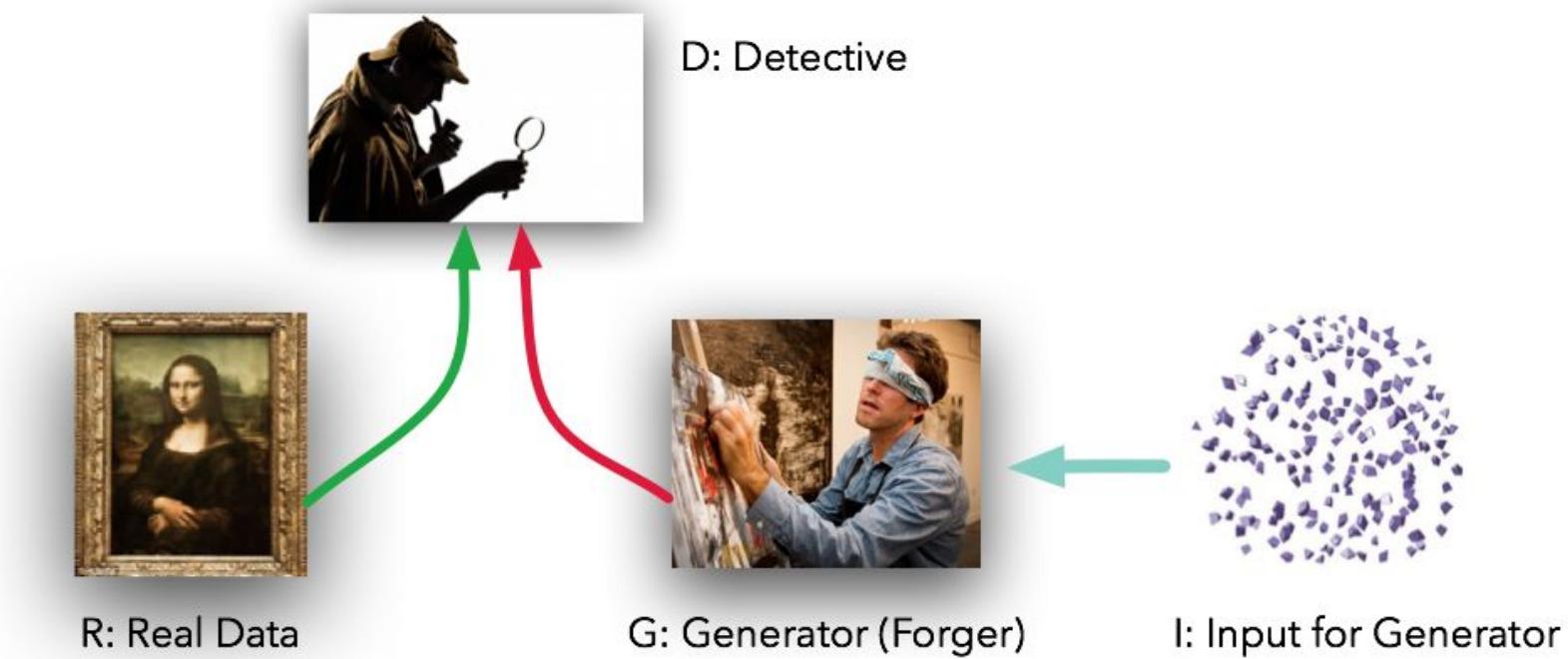


¿Cómo funciona? <https://arxiv.org/pdf/1406.2661.pdf>

- No es supervisado - no hay labels
- Vemos muchas imágenes y pedimos que genere imágenes de la misma distribución
- Discriminator dirige al generator

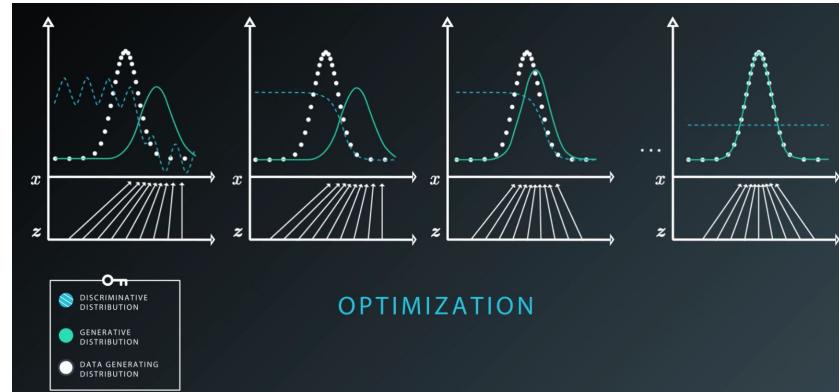


Discriminador

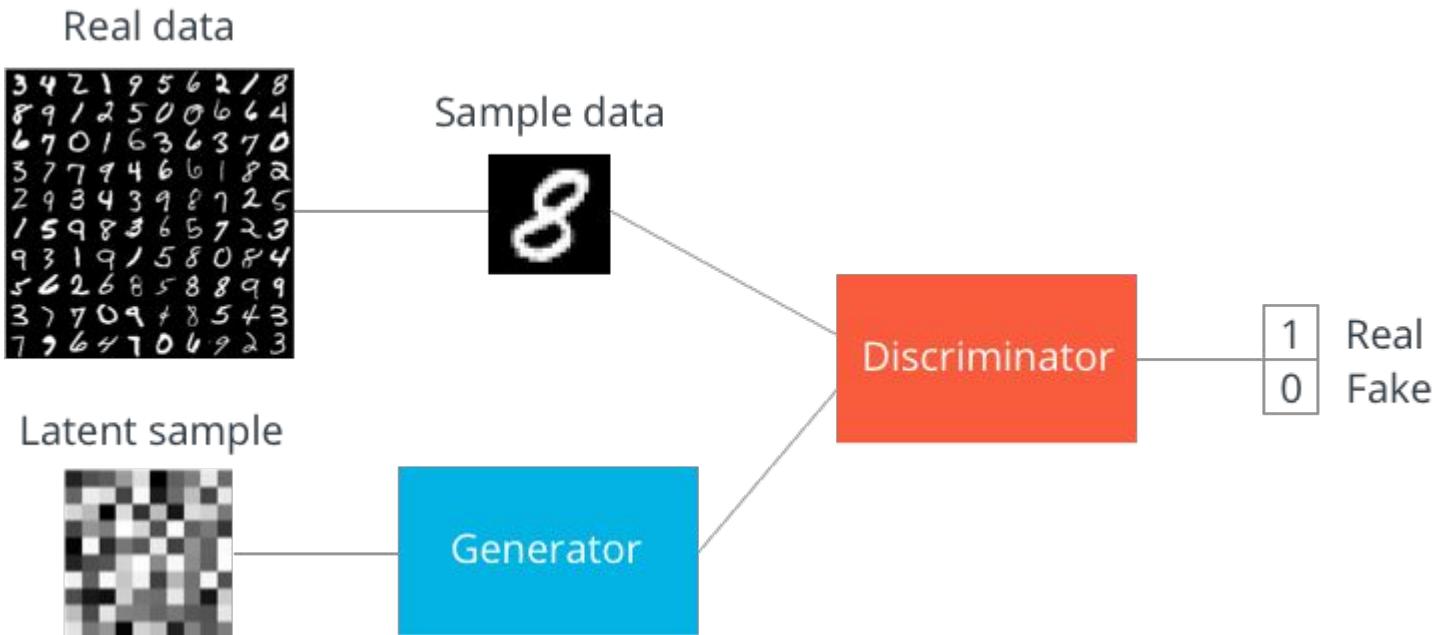


Adversarial

- Dos modelos compiten
- Utiliza teoría de juego (John von Neumann y John Nash)
 - Cooperación y conflicto entre agentes racionales
- Adaptamos la estrategia en el juego
- En un punto se llega a un **equilibrio**
 - Ningún jugador puede mejorar cambiando su estrategia



Concluyendo

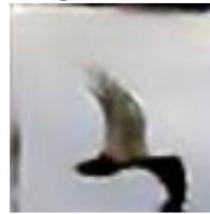


Text to Image StackGAN (2016)

This bird is white with some black on its head and wings, and has a long orange beak

This bird has a yellow belly and tarsus, grey back, wings, and brown throat, nape with a black face

This flower has overlapping pink pointed petals surrounding a ring of short yellow filaments



StackGAN
Stage-I
64x64
images

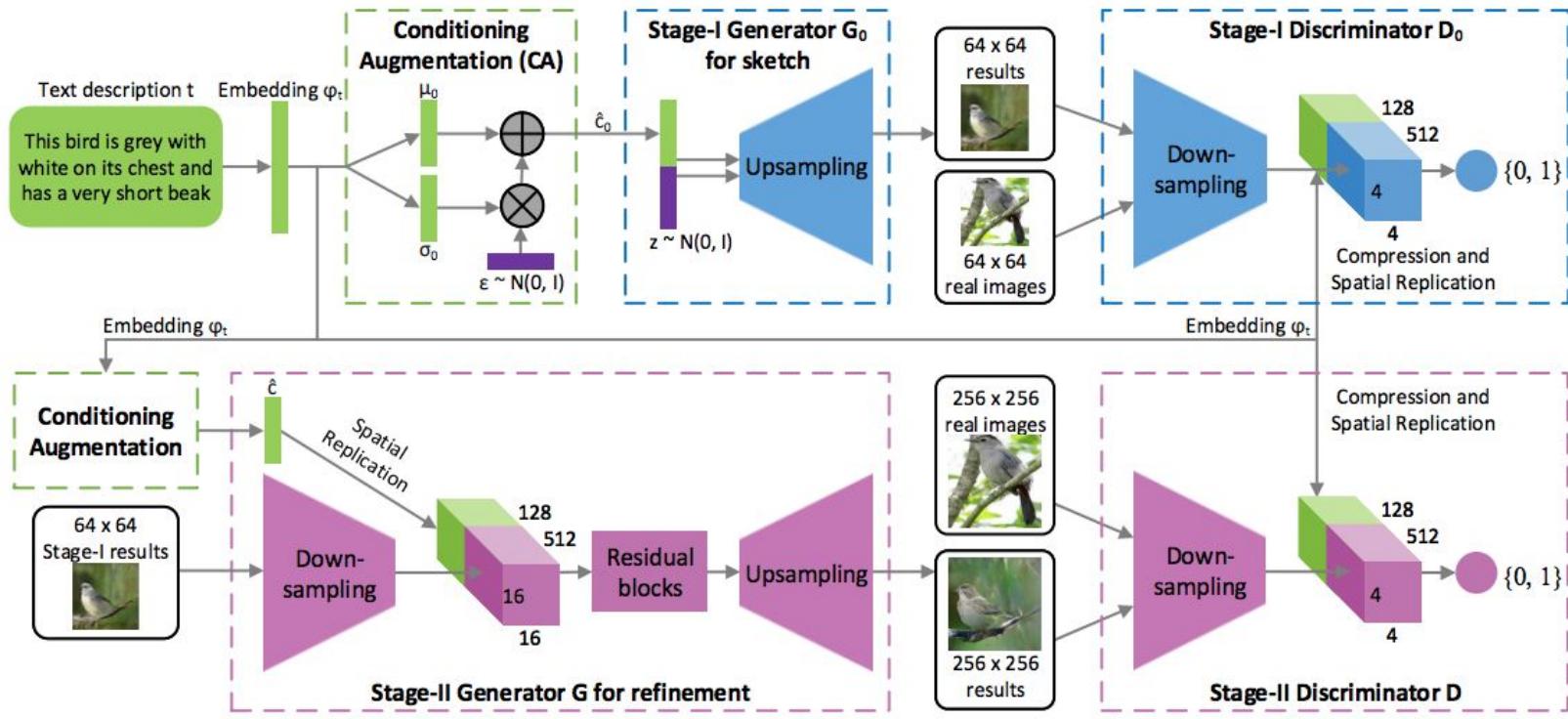
StackGAN
Stage-II
256x256
images



Vanilla GAN
256x256
images



Stage 1 GAN y Stage 2 GAN



A small yellow bird with a black crown and a short black pointed beak

Stage-I



Stage-II



Datasets

- CUB - 200 especies de pájaros con 11k imágenes , 10 descripciones
- Oxford-102 - 8200 flores, 10 descripciones
- MS-COCO - 120K imágenes variadas, cada una con 5 descripciones

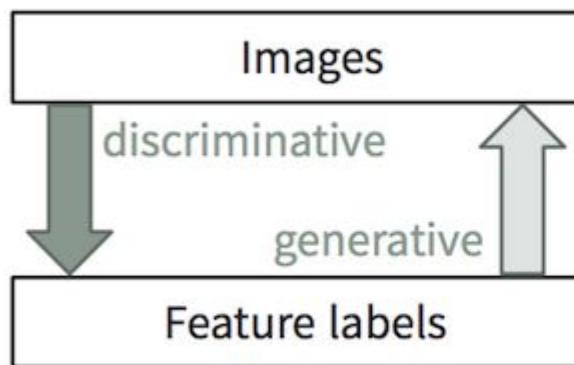
Text description	This flower has a lot of small purple petals in a dome-like configuration	This flower is pink, white, and yellow in color, and has petals that are striped	This flower has petals that are dark pink with white edges and pink stamen	This flower is white and yellow in color, with petals that are wavy and smooth	A picture of a very clean living room	A group of people on skis stand in the snow	Eggs fruit candy nuts and meat served on white dish	A street sign on a stoplight pole in the middle of a day
64x64 GAN-INT-CLS								
256x256 StackGAN								

Transparent Latent Space GAN (TL-GAN) (Oct 2018)

- [Demo](#)

Transparent Latent Space GAN (TL-GAN) (Oct 2018)

- ¿Podemos crear imágenes a partir de los labels?



Transparent Latent Space GAN (TL-GAN) (Oct 2018)

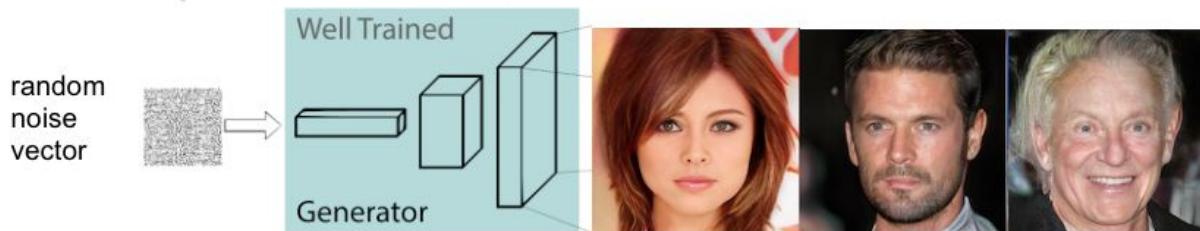
- ¿Podemos crear imágenes a partir de los labels?
 - pg-GAN de NVIDIA usa ruido de entrada



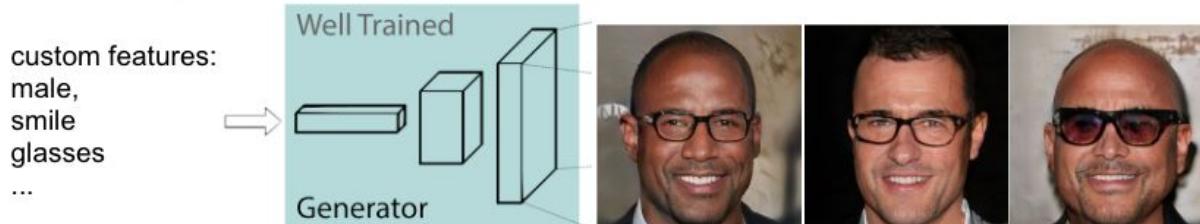
Transparent Latent Space GAN (TL-GAN) (Oct 2018)

- ¿Podemos crear imágenes a partir de los labels?

Random generation of high quality images



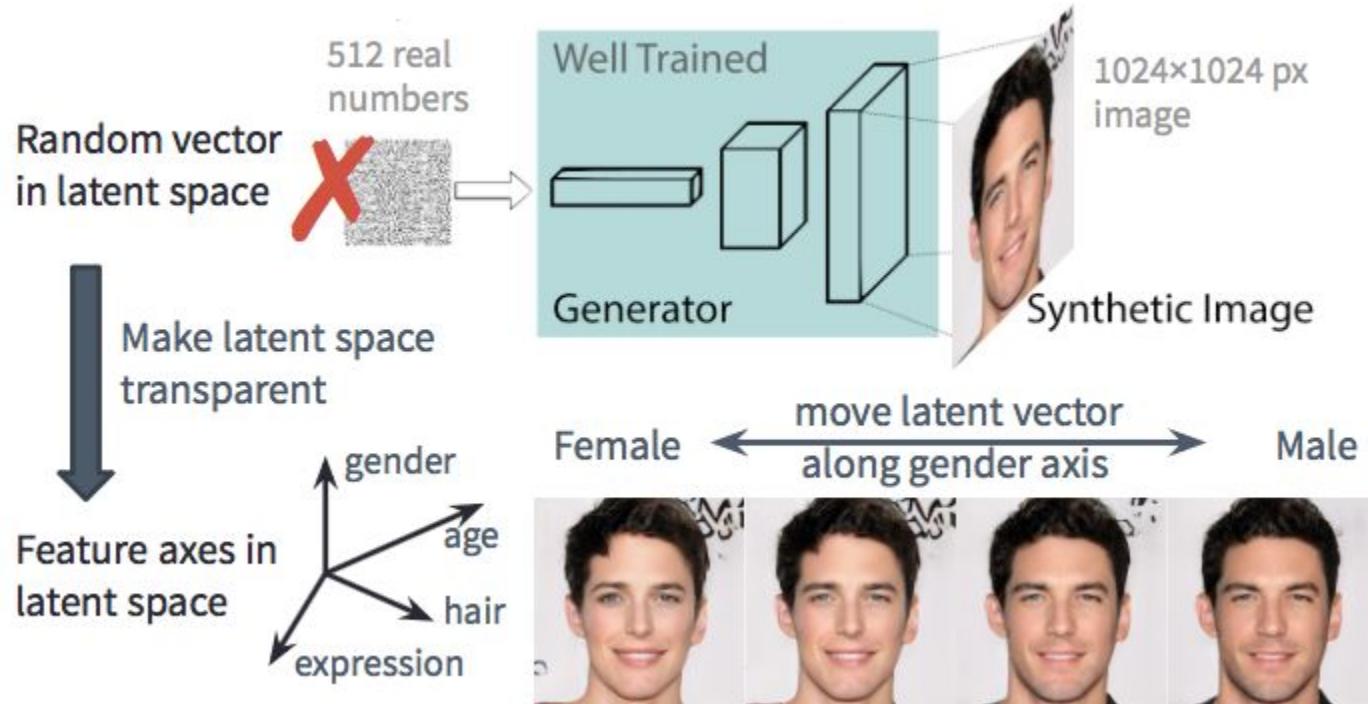
Controlled image generation according to custom features



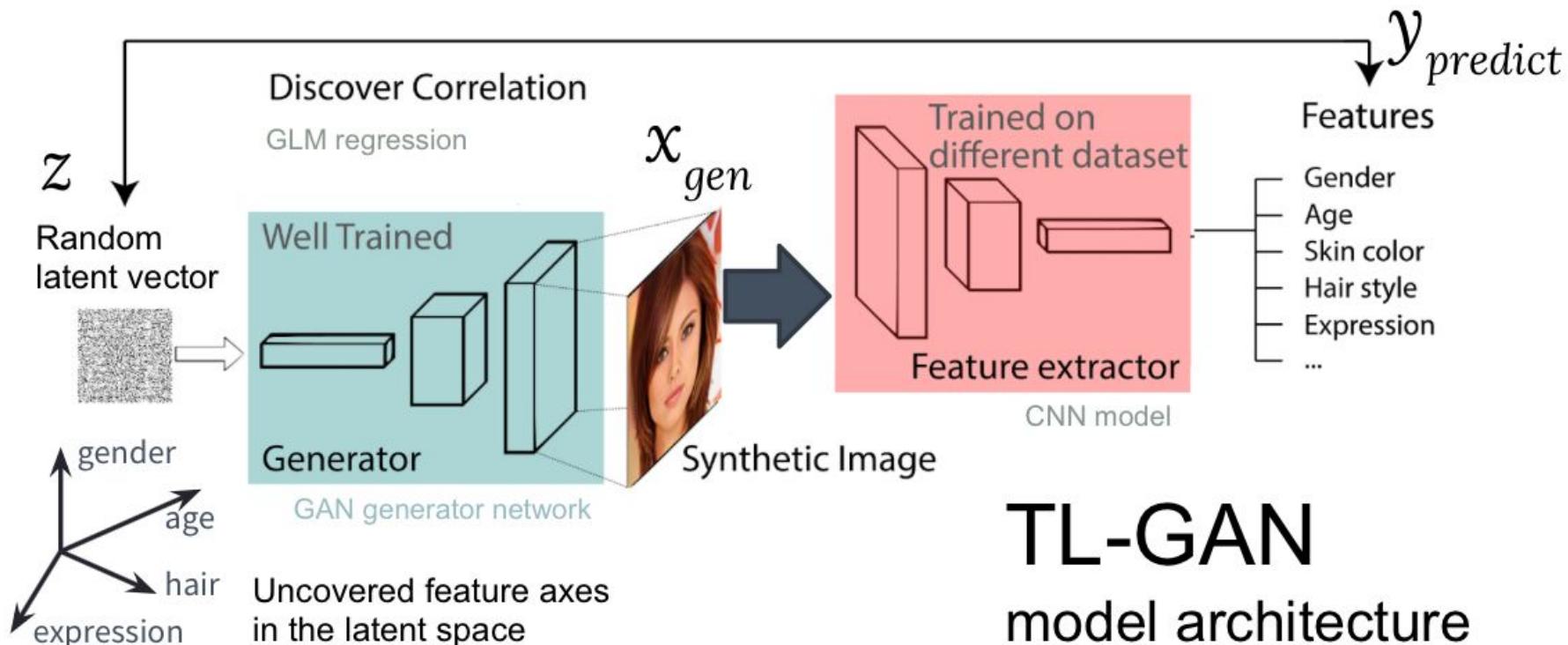
Transparent Latent Space GAN (TL-GAN) (Oct 2018)

- ¿Podemos crear imágenes a partir de los labels?
 - Style-transfer networks (pix2pix, CycleGAN) - convierten un dominio a otro
 - Generadores condicionales (Stack-GAN) - se debe volver a entrenar todo con nuevos features
- TL-GAN permite entrenar gradualmente la red
 - Se agregan features en menos de una hora

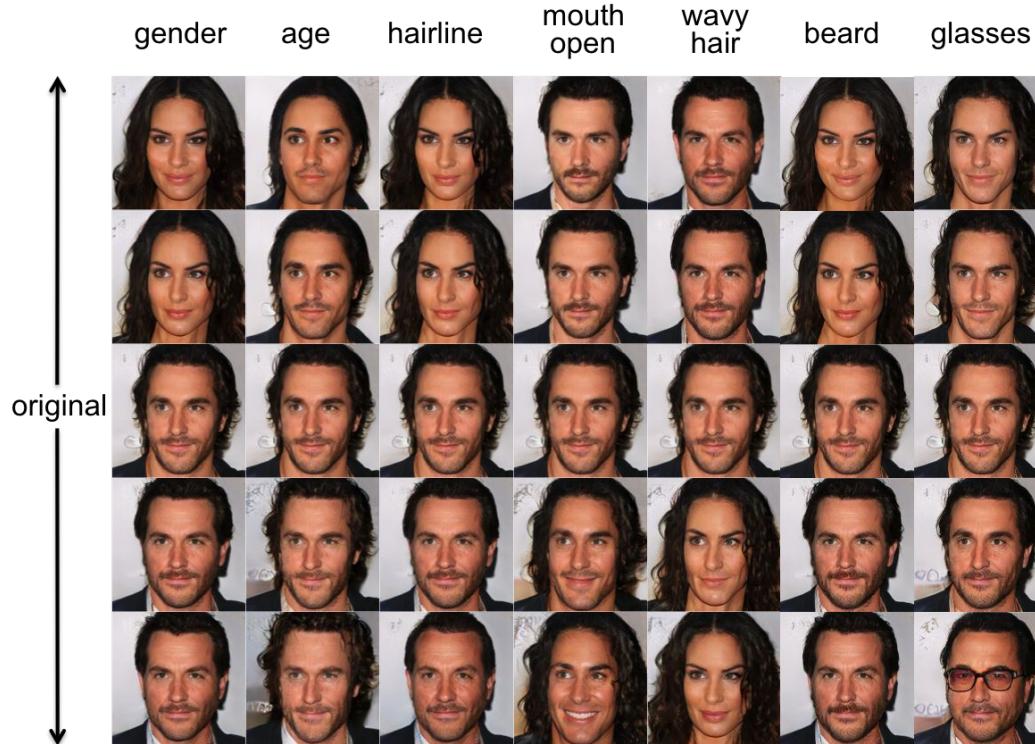
Transparent Latent Space GAN (TL-GAN) (Oct 2018)



Transparent Latent Space GAN (TL-GAN) (Oct 2018)



Transparent Latent Space GAN (TL-GAN) (Oct 2018)



Transparent Latent Space GAN (TL-GAN) (Oct 2018)



Big GAN (Google, Sept 2018)

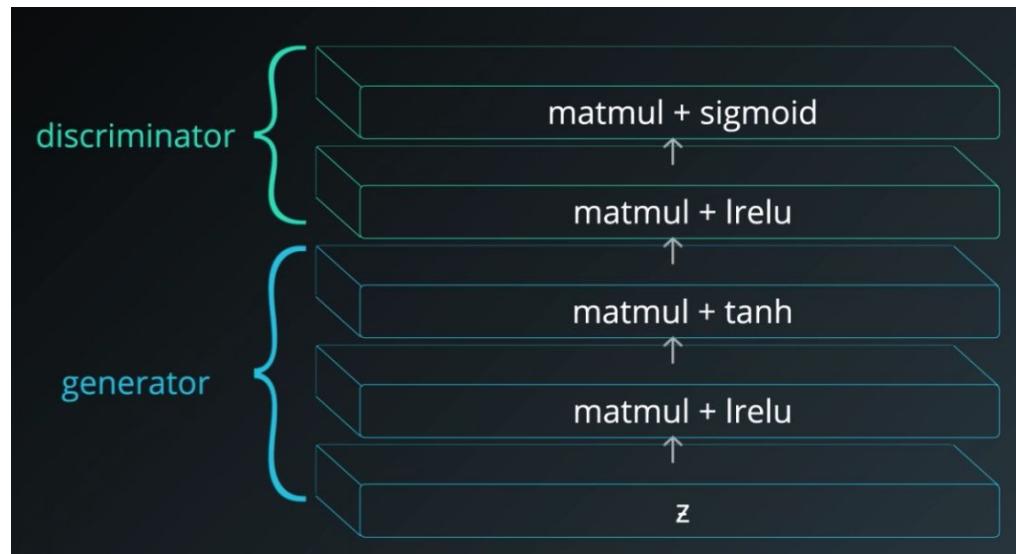


Gracias

- Image-to-Image Translation with Conditional Adversarial Networks
(<https://arxiv.org/abs/1611.07004>)
- Generative Visual Manipulation on the Natural Image Manifold
(<https://arxiv.org/abs/1609.03552>)
- Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks
(<https://arxiv.org/pdf/1703.10593.pdf>)
- StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks
<https://arxiv.org/abs/1612.03242>
- Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio "Generative Adversarial Networks", in NIPS 2014.
- Large Scale GAN Training for High Fidelity Natural Image Synthesis
(<https://arxiv.org/pdf/1809.11096.pdf>)
-

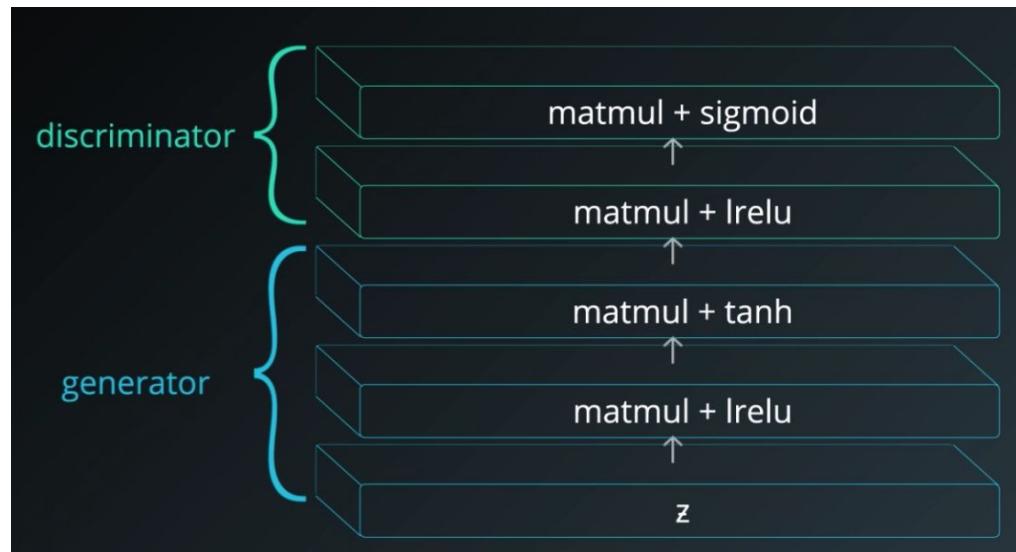
Arquitectura GAN

- Por lo menos una unidad oculta en discriminador y otra en generador
- Leaky RELUs funcionan muy bien en input de discriminador
- Tangente hiperbólica funciona bien como output del generador
- Output del discriminador suele ser probabilidad - sigmoid aplica



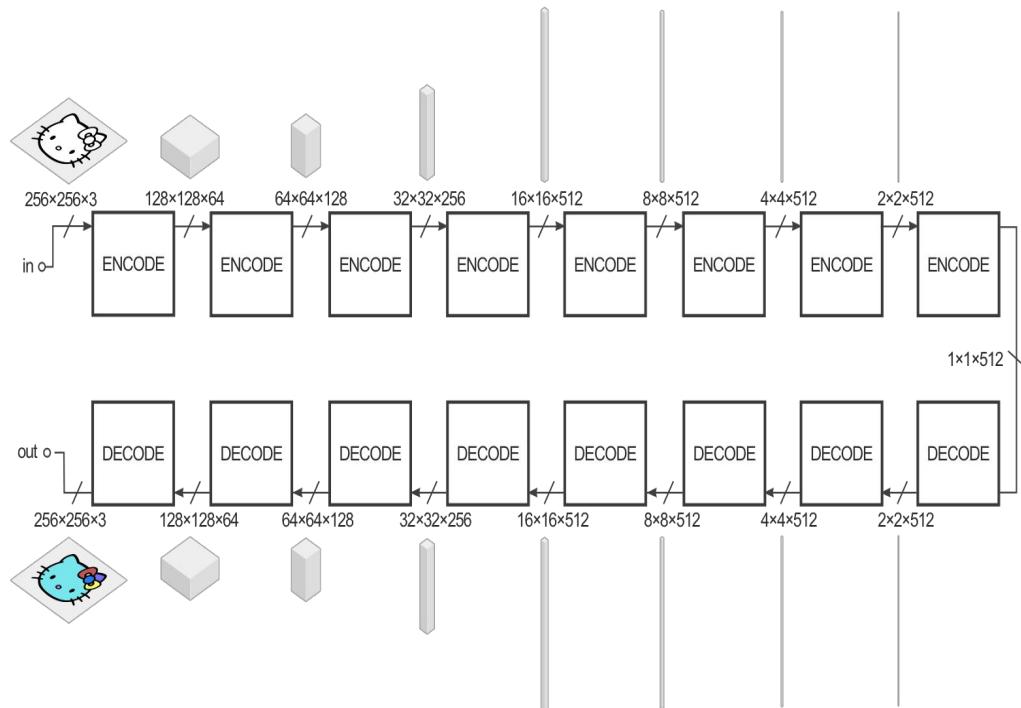
Arquitectura GAN

- Corremos dos optimizaciones simultáneamente
- Adam es bueno
- El discriminador funciona como clasificador normal
 - 1 -> REAL
 - 0 -> FALSO
- Función de pérdida
 - Cross entropy es buena idea
 - Volteado para generador



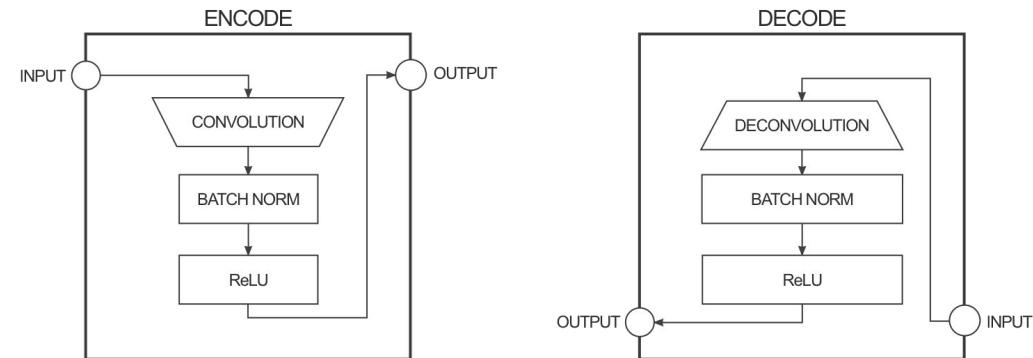
Generador de pix2pix (Berkeley, 2017) es un encoder-decoder

- Primero se hace el encoding
 - convoluciones + activaciones
 - Reduce la representación
 - Es una representación de nivel más alto



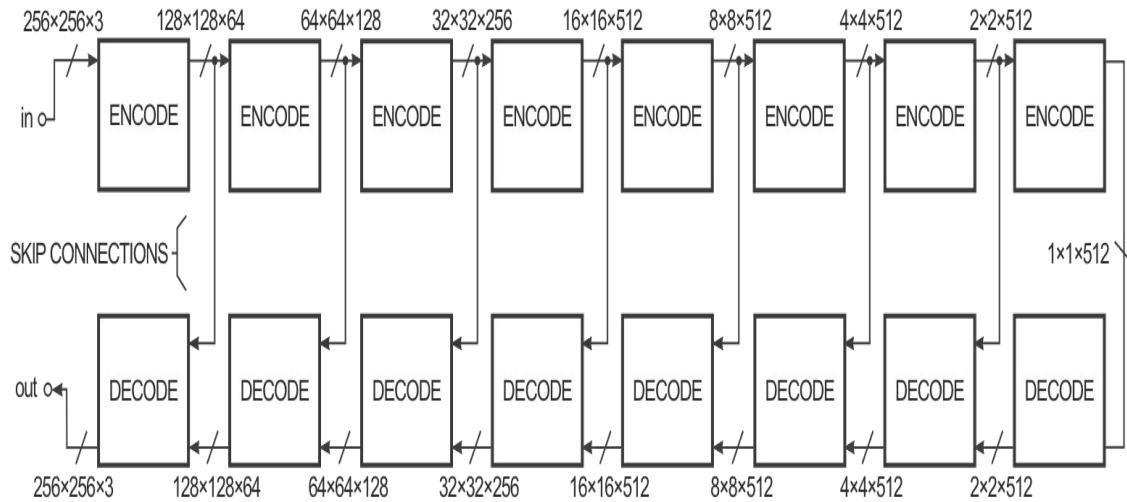
Generador de pix2pix (Berkeley, 2017) es un encoder-decoder

- Primero se hace el encoding
 - convoluciones + activaciones
 - Reduce la representación
 - Es una representación de nivel más alto



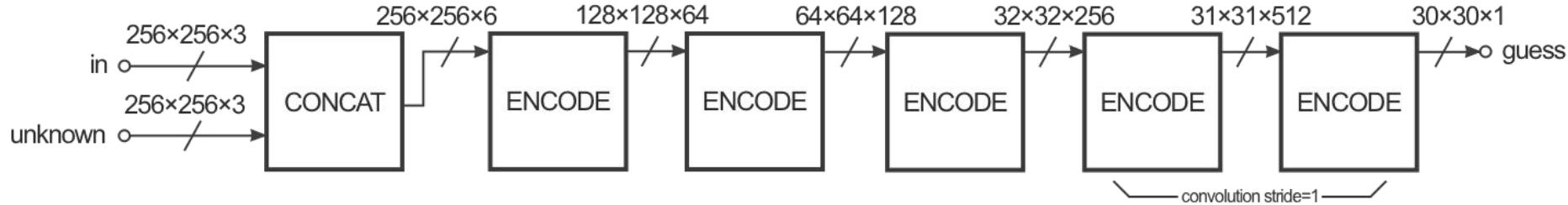
Generador de pix2pix (Berkeley, 2017) es un encoder-decoder

- Primero se hace el encoding
 - convoluciones + activaciones
 - Reduce la representación
 - Es una representación de nivel más alto
- Agregan skip connections para saltar los encoders-decoders si no son necesarios



Discriminador pix2pix (Berkeley, 2017)

- Recibe dos imágenes
 - Una de entrada y una desconocida
- Determina si la segunda fue creada por el generador
- La salida es una imagen de 30x30
 - Cada pixel tiene valor de 0 a 1
 - Cada pixel representa qué tan confiable es una sección de la imagen
 - Cada pixel de la salida corresponde a un patch de 70x70 de la original (**PatchGAN**)



Entrenamiento de pix2pix

