

Generative Adversarial Networks

Israel N. Chaparro-Cruz

Presentación



- MSc.c en Ciencia de la Computación - UCSP - Beca FONDECYT - CONCYTEC.
- Candidato a MSc. en Sistemas e Informática ATI - UNJBG.
- Ingeniero en Informática y Sistemas - UNJBG.
- Docente del DAIS - UNJBG.
- Poster Presenter en Uruguay, Canadá y USA.
- Autor del paper:
 - BORDE: Boundary and Sub-Region Brain Image Synthesis.

Motivación: ¿Cuál imagen es falsa?



<https://thispersondoesnotexist.com/>

<https://thispersondoesnotexist.com/>

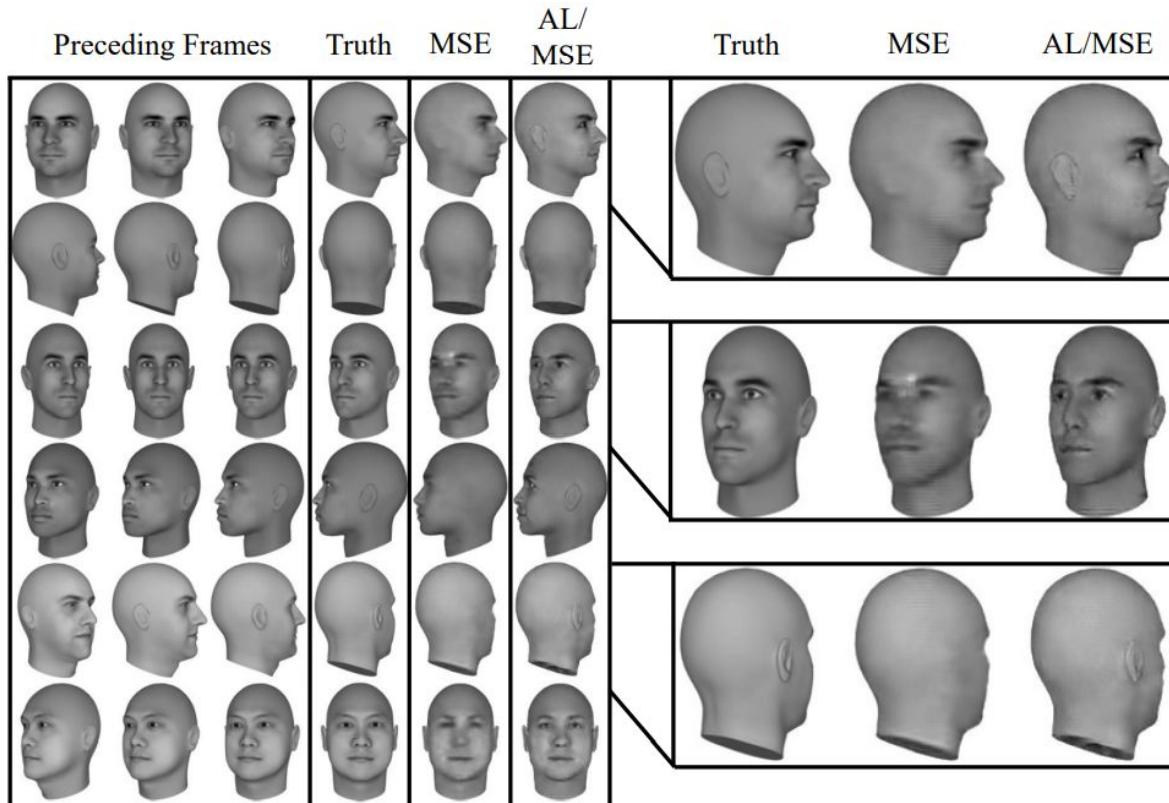
GANs

- Generativas
 - Aprenden un modelo generativo
- Adversarial
 - Entrenadas en una configuración adversarial
- Networks
 - Usan Redes Neuronales Profundas

¿Por qué Modelos Generativos?

- vs Discriminativos
 - Dada una imagen \mathbf{X} , predecir la clasificación \mathbf{Y} .
 - Estimar $\mathbf{P}(\mathbf{Y}|\mathbf{X})$.
- Los modelos discriminativos tienen varias limitaciones clave
 - No pueden modelar $\mathbf{P}(\mathbf{X})$, i.e. la probabilidad de ver una determinada imagen.
 - Por lo tanto, no pueden tomar una muestra de $\mathbf{P}(\mathbf{X})$, i.e. no pueden generar nuevas imágenes.
- Los modelos generativos (en general) hacen frente a todo lo anterior
 - Pueden modelar $\mathbf{P}(\mathbf{X})$.
 - Pueden generar nuevas imágenes.

Modelos Generativos vs Discriminativos



Lotter (2015). Unsupervised learning
of visual structure using predictive
generative networks.

GANs

- Generativas
 - Aprenden un modelo generativo
- Adversarial
 - Entrenadas en una configuración adversarial
- Networks
 - Usan Redes Neuronales Profundas

Entrenamiento Adversarial



+ 0.01 ×



=



“panda”

81.97% confidence

“papillon dog”

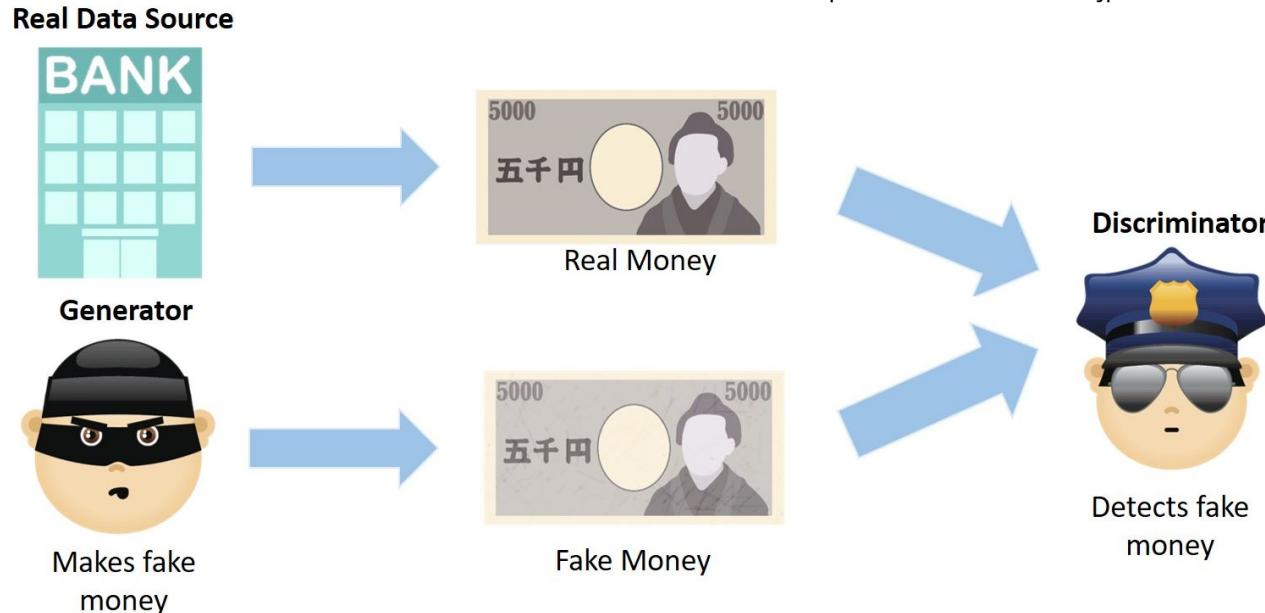
99.56% confidence

https://www.macnica.co.jp/business/ai_iot/columns/135130/

- Podemos generar muestras adversas para **engañar** al modelo discriminativo.
- Podemos usar esas muestras adversas para hacer los modelos más **robustos**.
- Entonces requerimos más esfuerzo para **generar** muestras adversas.
- Repetimos esto y obtenemos **mejores** modelos discriminativos.

Entrenamiento Adversarial

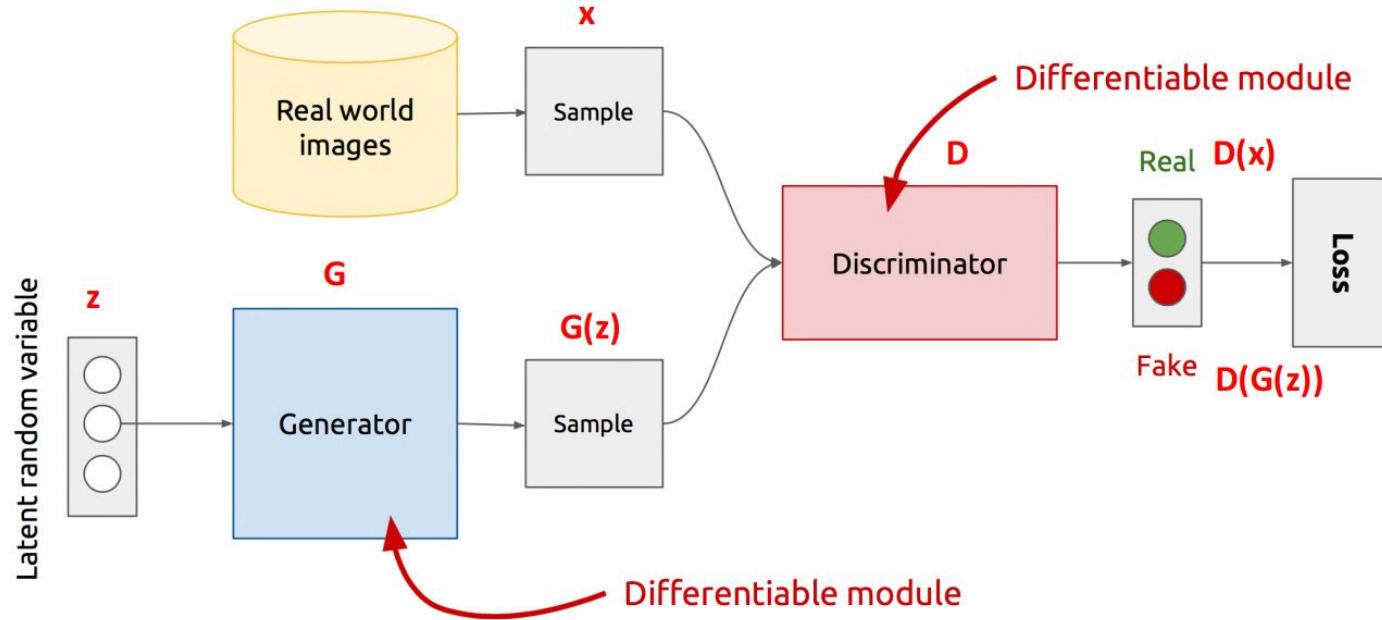
https://www.macnica.co.jp/business/ai_iot/columns/135130/



- Entrenamiento uno **contra** el otro.
- **Repetir** para obtener un mejor Generador y Discriminador

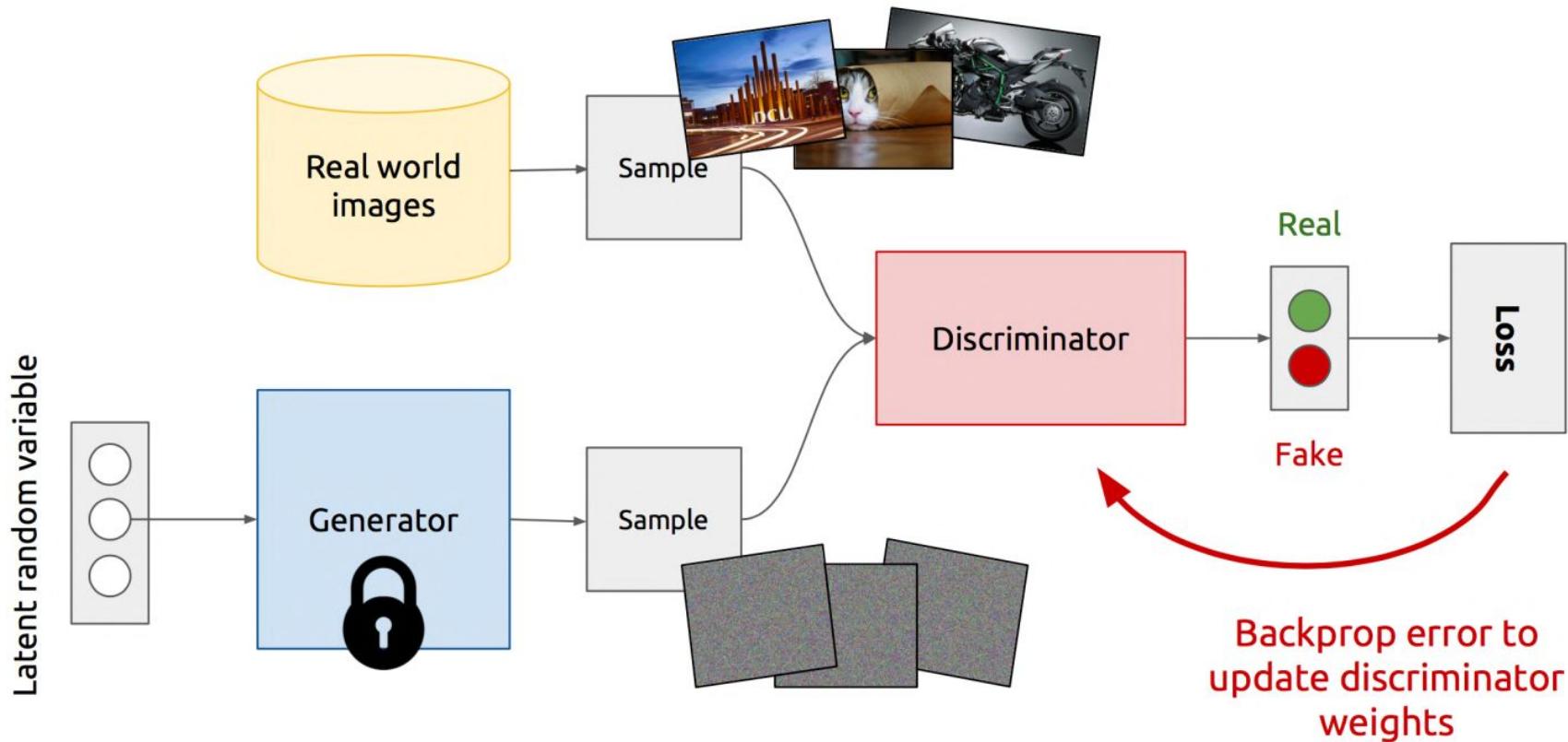
Arquitectura de una GANs

<https://www.slideshare.net/xavigiro/deep-learning-for-computer-vision-generative-models-and-adversarial-training-upc-2016>



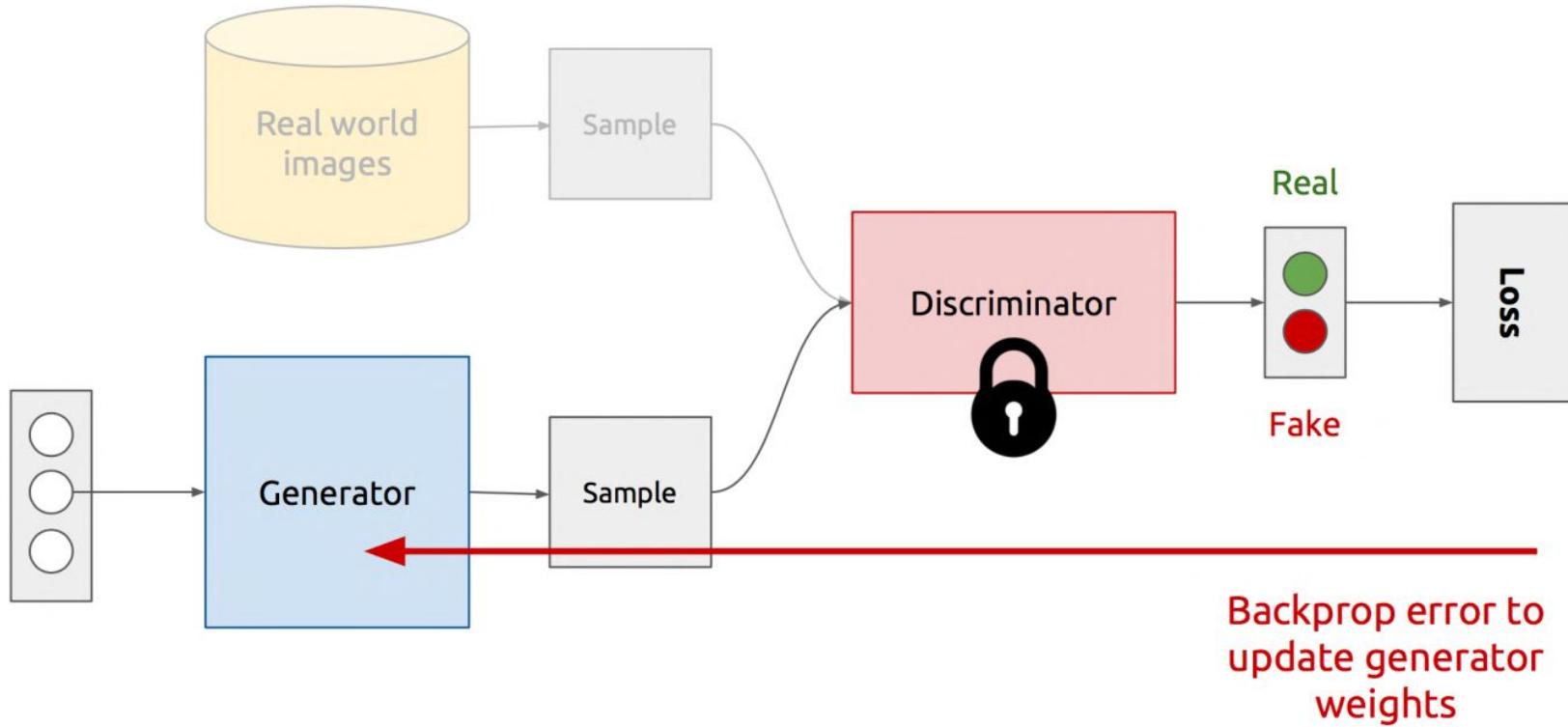
- Z es algún **ruido** aleatorio (gaussiano/uniforme).
- Z puede pensarse como la representación **latente** de la imagen.

Entrenando al Discriminador



Entrenando al Generador

Latent random variable

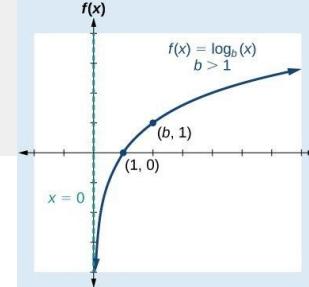


GANs

- Generativas
 - Aprenden un modelo generativo
- Adversarial
 - Entrenadas en una configuración adversarial
- Networks
 - Usan Redes Neuronales Profundas

Formulación de las GANs

$$\min_G \max_D V(D, G)$$



- Es formulado como un **minmax game**, donde:
 - El discriminador intenta **maximizar** la recompensa $V(D, G)$.
 - El generador intenta **minimizar** la recompensa del discriminador.

$$V(D, G) = \mathbb{E}_{x \sim p(x)} [\log D(x)] + \mathbb{E}_{z \sim q(z)} [\log(1 - D(G(z)))]$$

- El equilibrio de nash de este juego en particular se logra cuando:
 - $P_{data}(x) = P_{gen}(x) \quad \forall x$
 - $D(x) = \frac{1}{2} \quad \forall x$

**Discriminator
updates**

Algorithm 1 Minibatch stochastic gradient descent training of generative adversarial nets. The number of steps to apply to the discriminator, k , is a hyperparameter. We used $k = 1$, the least expensive option, in our experiments.

for number of training iterations **do**

for k steps **do**

- Sample minibatch of m noise samples $\{\mathbf{z}^{(1)}, \dots, \mathbf{z}^{(m)}\}$ from noise prior $p_g(\mathbf{z})$.
- Sample minibatch of m examples $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(m)}\}$ from data generating distribution $p_{\text{data}}(\mathbf{x})$.
- Update the discriminator by ascending its stochastic gradient:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m \left[\log D(\mathbf{x}^{(i)}) + \log (1 - D(G(\mathbf{z}^{(i)}))) \right].$$

end for

- Sample minibatch of m noise samples $\{\mathbf{z}^{(1)}, \dots, \mathbf{z}^{(m)}\}$ from noise prior $p_g(\mathbf{z})$.
- Update the generator by descending its stochastic gradient:

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log (1 - D(G(\mathbf{z}^{(i)}))).$$

**Generator
updates**

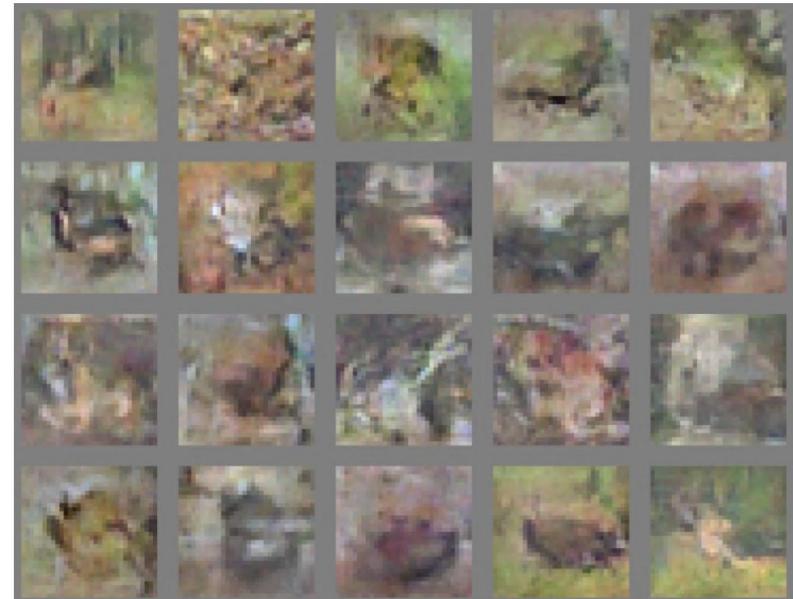
end for

The gradient-based updates can use any standard gradient-based learning rule. We used momentum in our experiments.

Faces



CIFAR



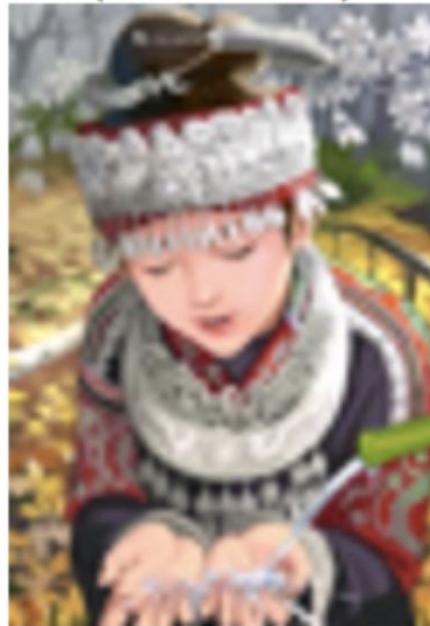
Goodfellow (2014), Generative Adversarial Networks.

Magia de las GANs

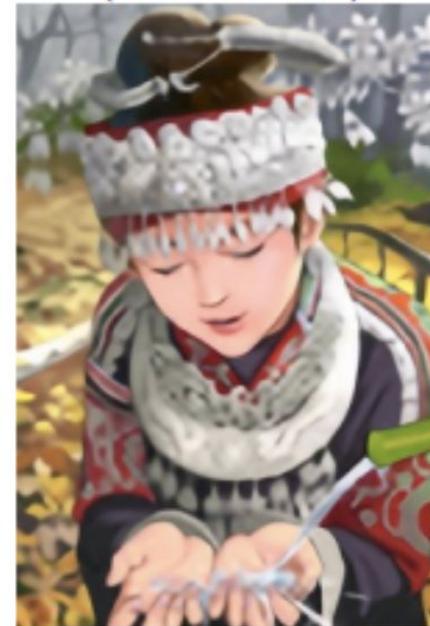
original



bicubic
(21.59dB/0.6423)



SRResNet
(23.44dB/0.7777)



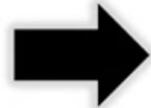
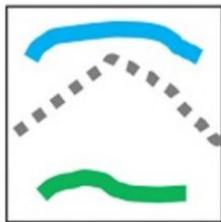
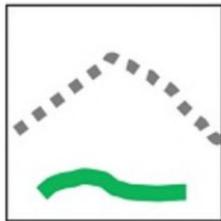
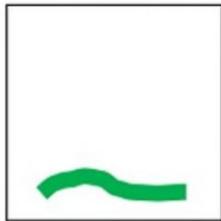
SRGAN
(20.34dB/0.6562)



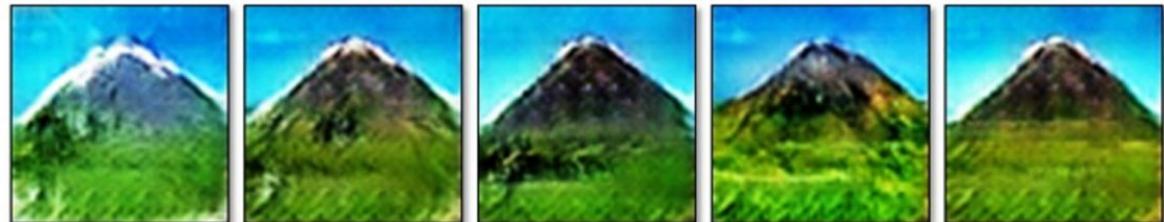
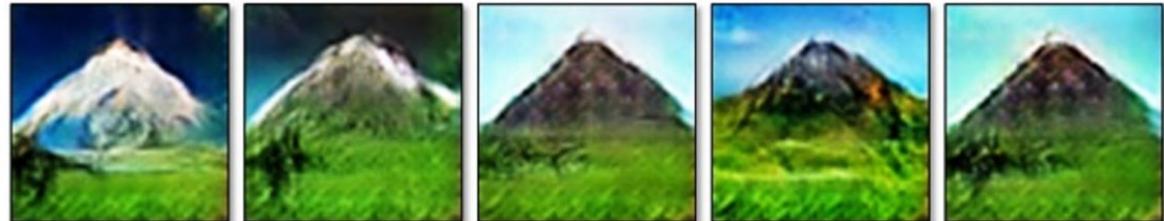
<https://arxiv.org/pdf/1701.00160.pdf>

Magia de las GANs

User edits

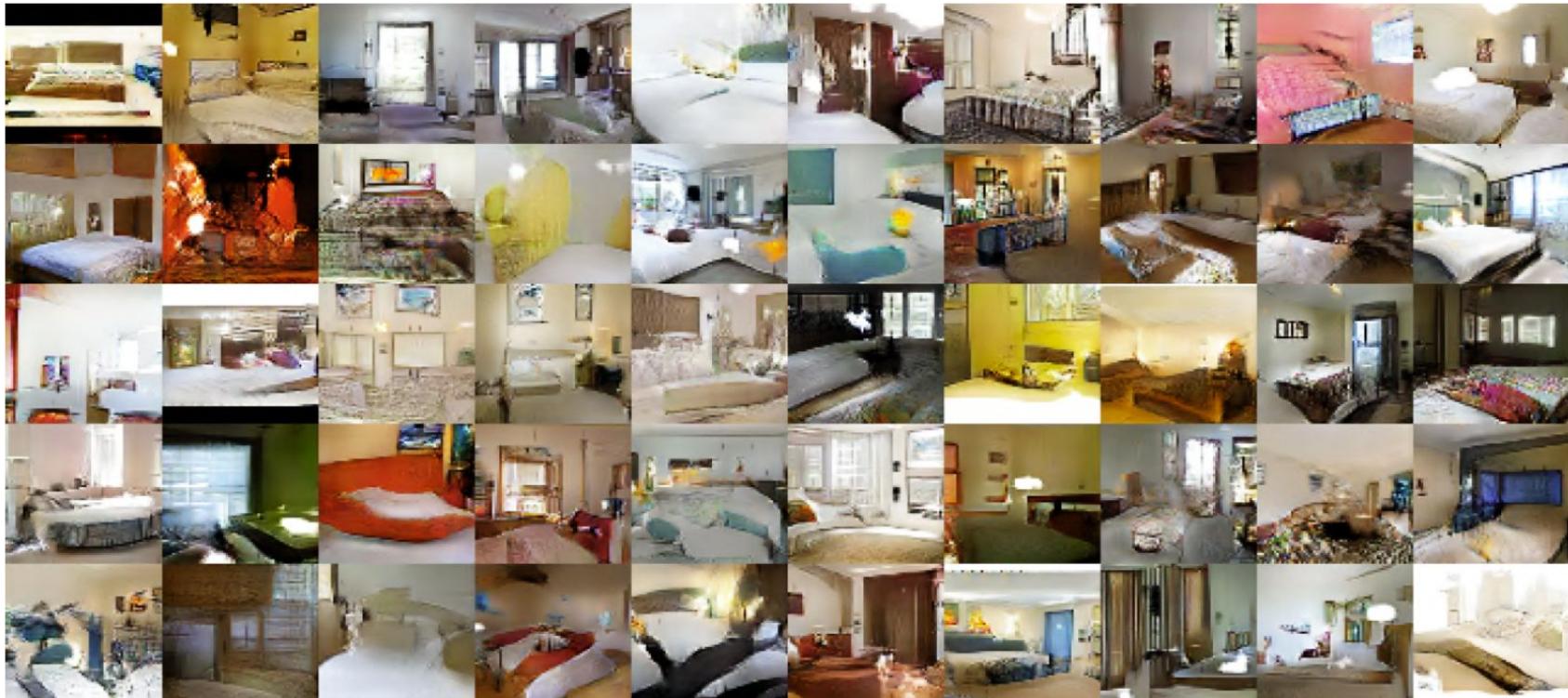


Generated images



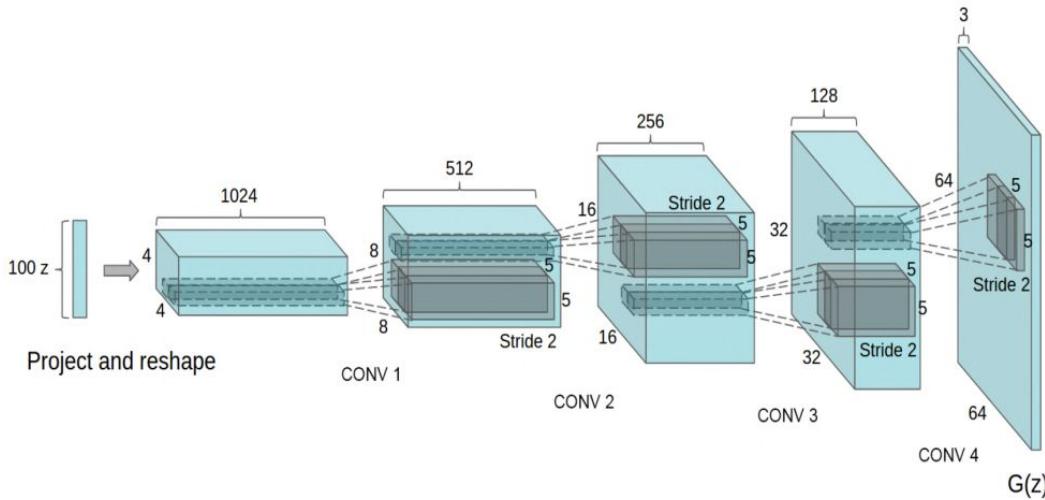
<http://people.eecs.berkeley.edu/~junyanz/projects/gvm/>

DCGAN: Bedroom images



Deep Convolutional GANs (DCGANs)

Generator Architecture

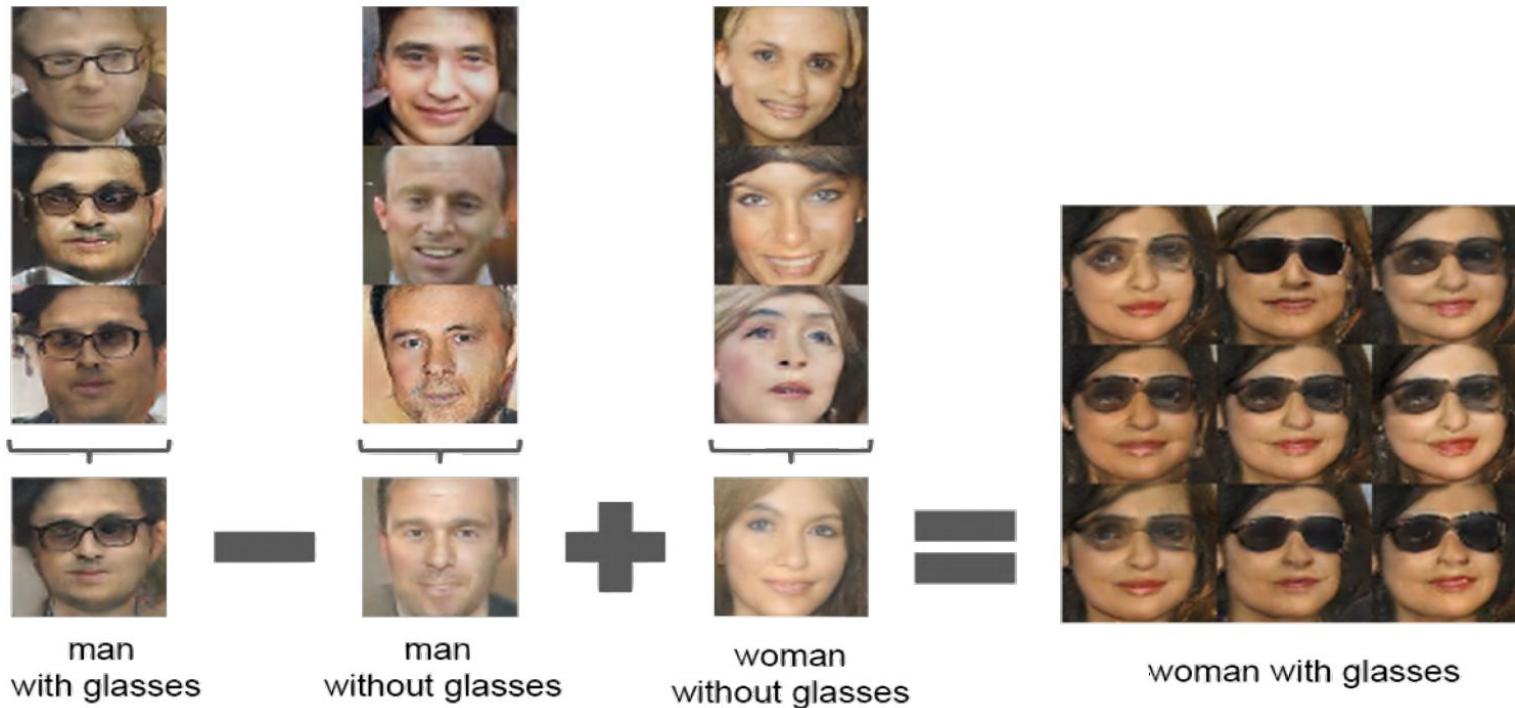


Ideas clave:

- Reemplazar las capas ocultas totalmente conectadas con convoluciones.
 - **Generador:** Fractional-Strided convolutions.
- Usar batch-normalization después de cada capa.
- **Dentro del Generador:**
 - Usar ReLU para las capas ocultas.
 - Usar Tanh para la capa de salida.

Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." arXiv:1511.06434 (2015).

Aritmética de vectores latentes



Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." arXiv:1511.06434 (2015).

Ventajas de las GANs

- Existencia plena de trabajos en Modelos Generativos Profundos
 - Boltzmann Machine
 - Deep Belief Nets
 - Variational AutoEncoders (VAE)
- ¿Por qué GANs?
 - Muestreo (generación) es sencilla.
 - El entrenamiento no involucra estimación de máxima verosimilitud.

Problemas con las GANs

- La distribución de probabilidad es implícita
 - No es sencillo calcular $P(X)$.
 - Por lo tanto, los Vanilla GANs solo son buenos para muestreo/generación.
- El entrenamiento es difícil
 - No-convergencia
 - Modo colapso

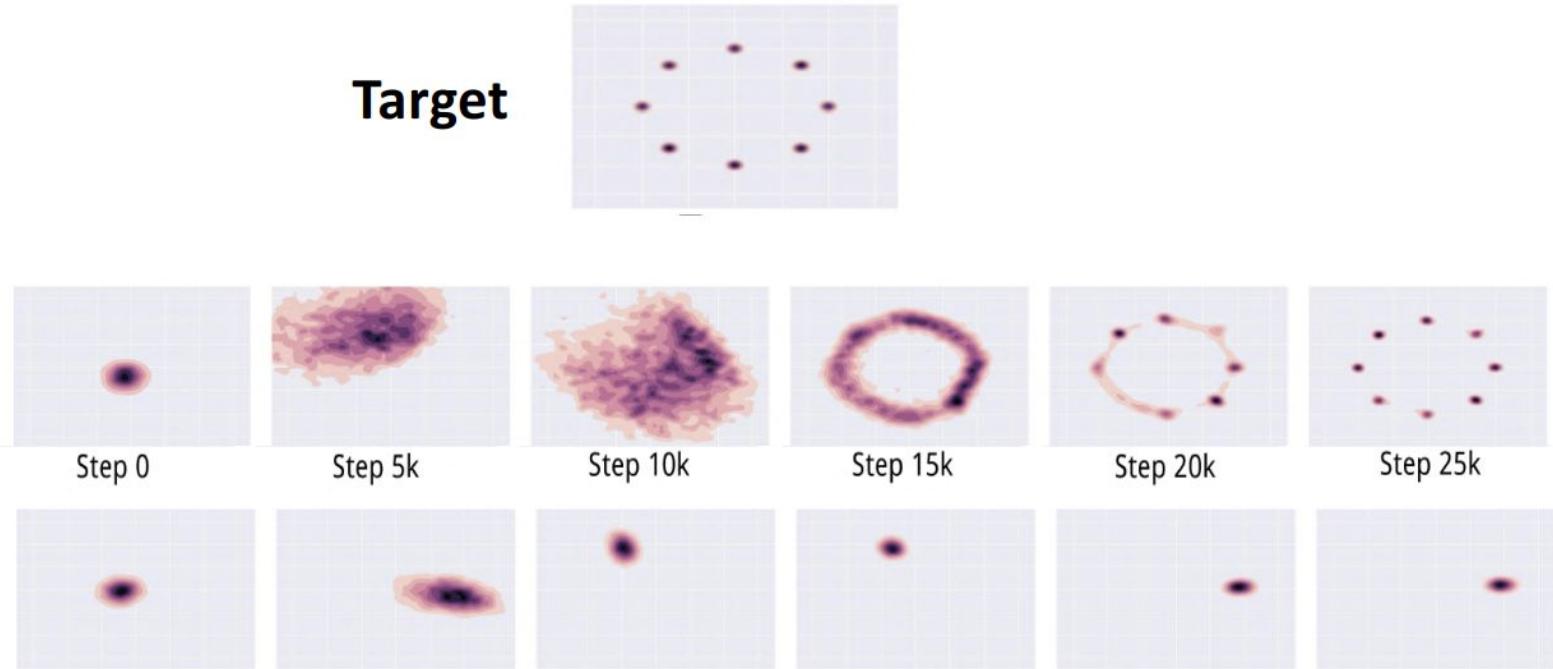
No-convergencia

$$\min_G L(G)$$

$$\min_G \max_D V(D, G)$$

- Los modelos de deep-learning en general tienen **un solo jugador**
 - El jugador trata de **maximizar** su recompensa (minimizar el loss).
 - Usa SGD (con **Backpropagation**) para encontrar los parámetros óptimos.
 - SGD garantiza la **convergencia** (bajo ciertas condiciones).
 - Problema: Con no-convexidad, podríamos converger a óptimos **locales**.
- Las GANs en cambio involucra dos (o más) jugadores
 - El discriminador está tratando de **maximizar** su recompensa.
 - El generador está tratando de **minimizar** la recompensa del generador.
 - **Problema:** Podríamos no converger en absoluto al equilibrio de Nash.

Modo-colapso



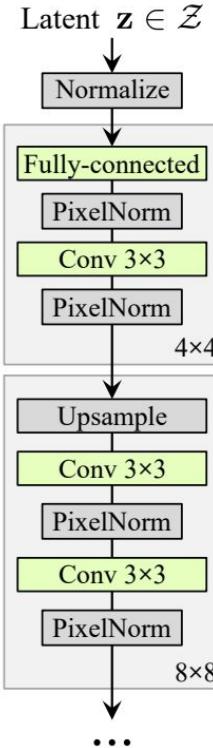
Metz, Luke, et al. "Unrolled Generative Adversarial Networks." arXiv preprint arXiv:1611.02163 (2016).

Modo-colapso

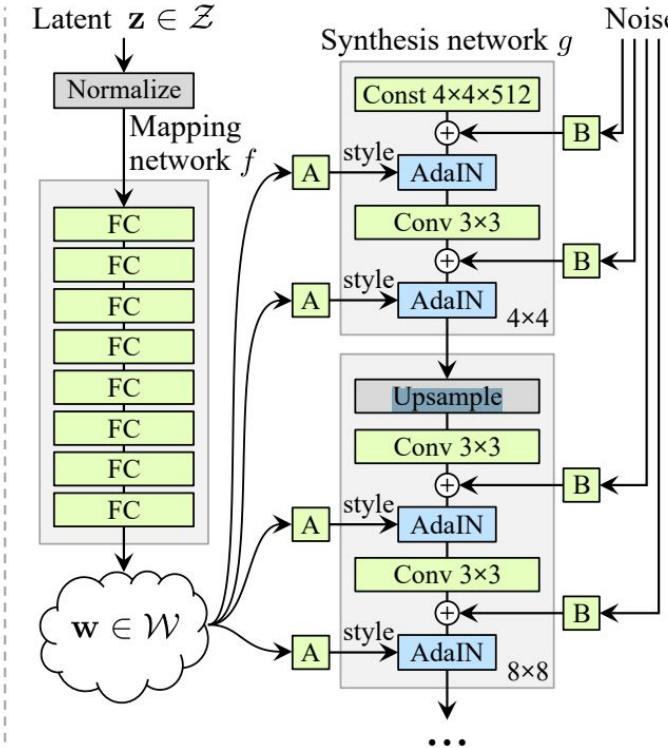


Reed, S., et al. *Generating interpretable images with controllable structure*. Technical report, 2016. 2, 2016.

StyleGAN



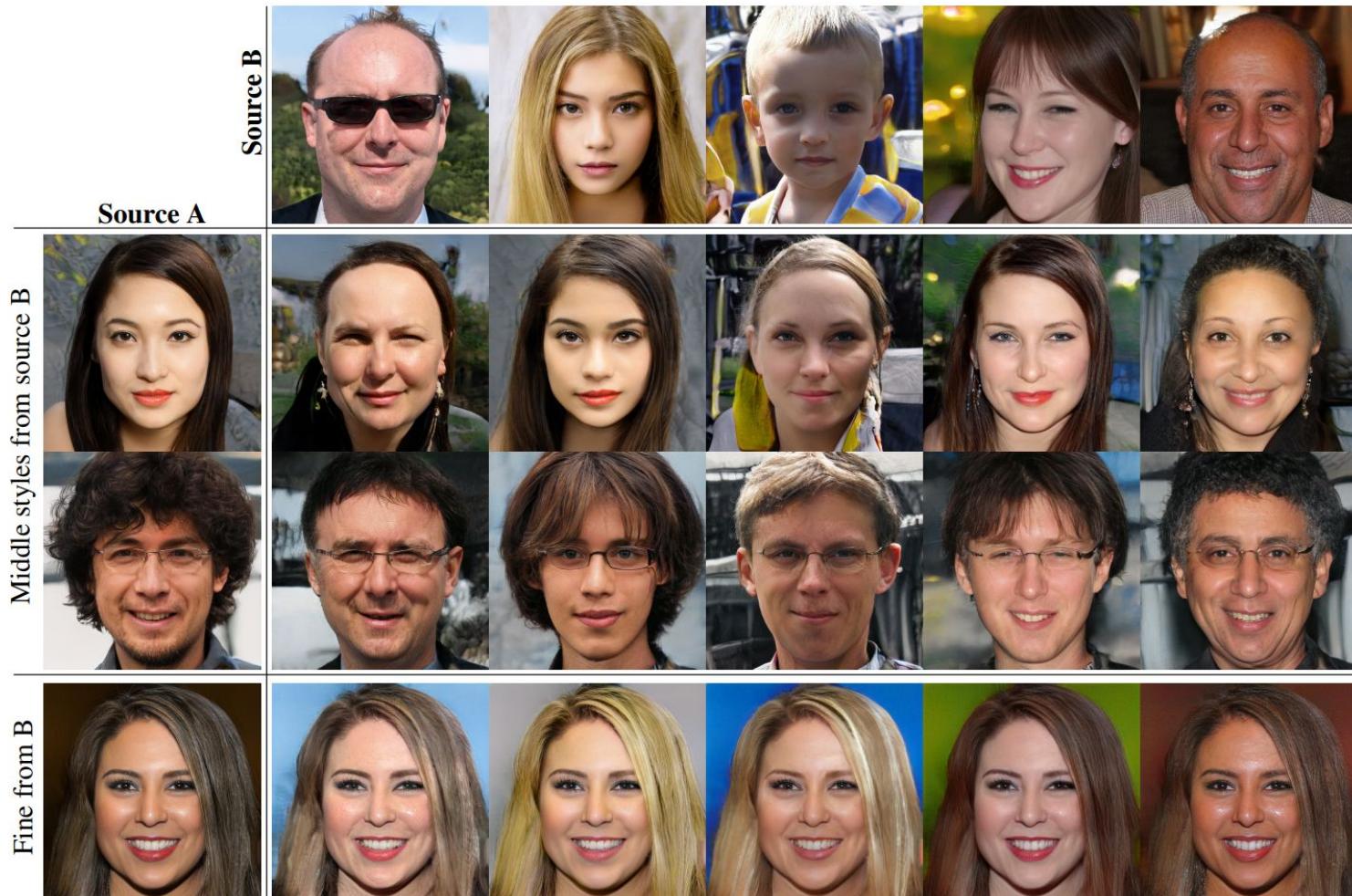
(a) Traditional



(b) Style-based generator

Karras, T. (2019). StyleGAN





Input

Generation from a single Image



Input 1

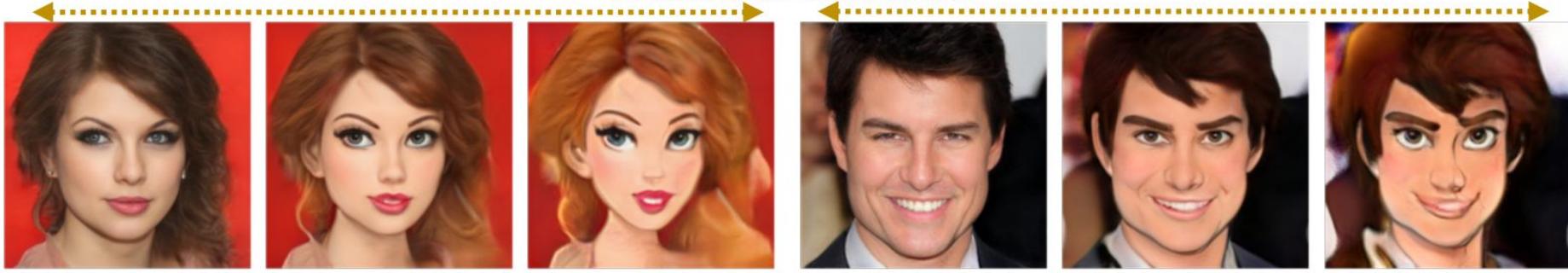
Input 2

Feature Interpolation



Chong (2021). StyleGAN of All Trades

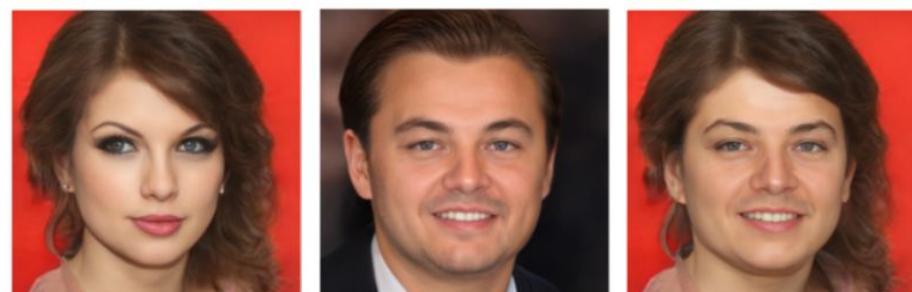
Continuous Translation



Local Image Translation



Attributes Transfer



Chong (2021). StyleGAN of All Trades

GANs condicionadas

MNIST dígitos generados condicionados a su etiqueta de clase.

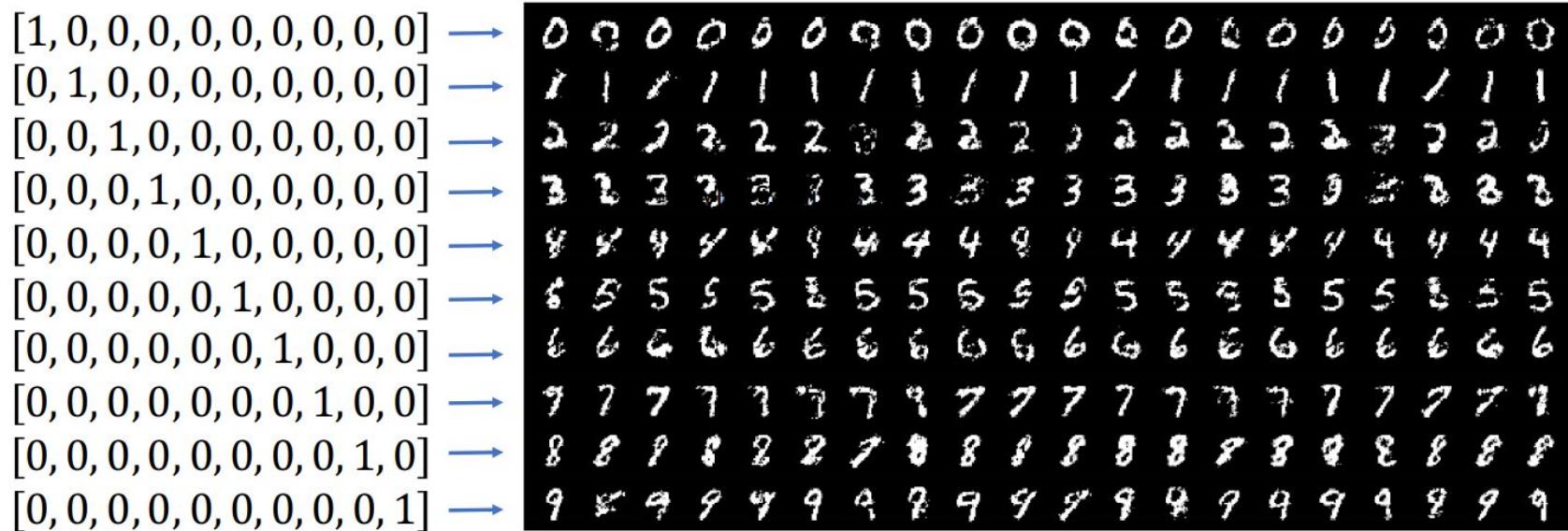
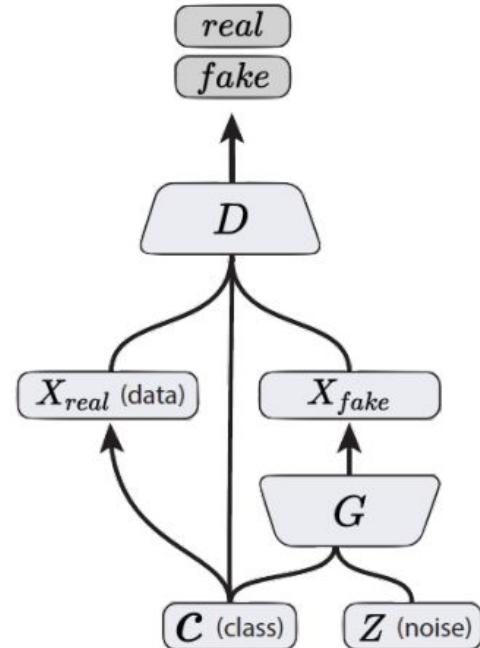


Figure 2 in the original paper.

GANs condicionadas

- Modificación simple al marco GAN original que condiciona el modelo a información adicional para un mejor aprendizaje multimodal.
- Se presta a muchas aplicaciones prácticas de GAN cuando tenemos supervisión explícita disponible.



Conditional GAN
(Mirza & Osindero, 2014)

Aplicación: Traducción de imagen a imagen

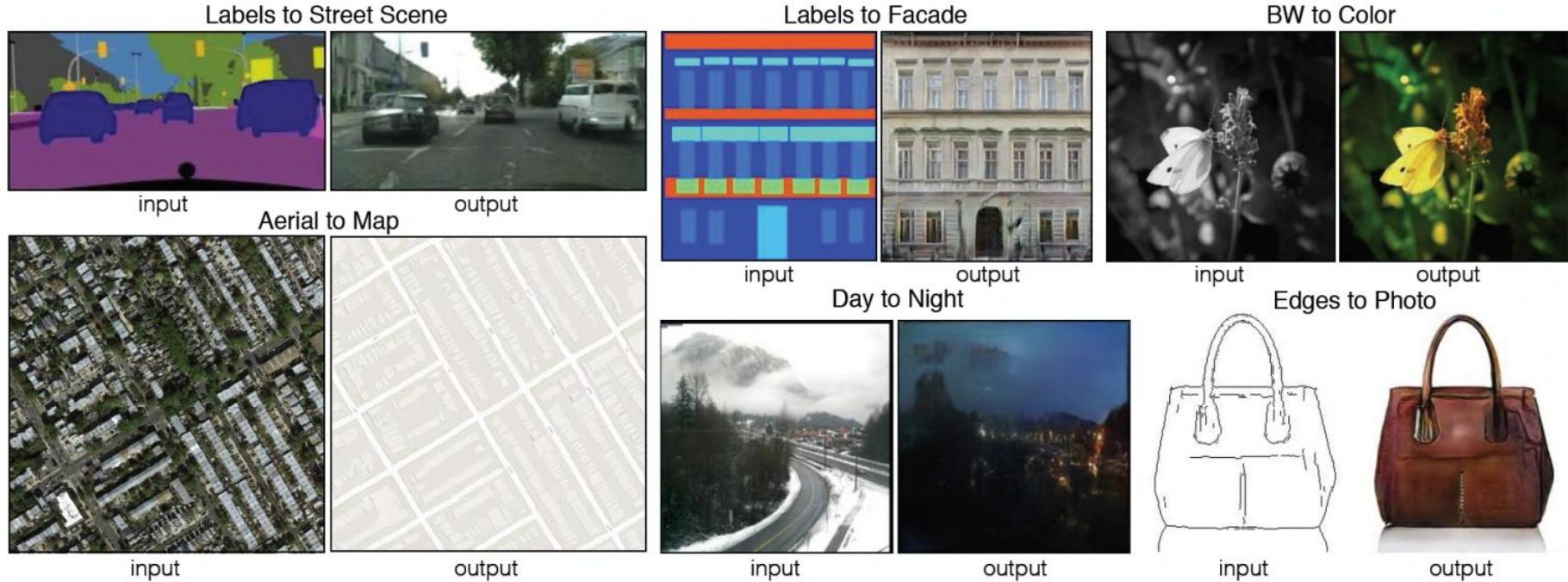


Figure 1 in the original paper.

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. "Image-to-image translation with conditional adversarial networks". arXiv preprint arXiv:1611.07004. (2016).

Aplicación: Traducción de imagen a imagen

- Arquitectura: basado en la DCGAN.
- El entrenamiento está condicionado en las imágenes desde el dominio fuente.
- Modificación simple al marco GAN original que condiciona el modelo a información adicional para un mejor aprendizaje multimodal.
- Las GANs condicionales brindan una forma efectiva de manejar muchos dominios complejos sin preocuparse por diseñar funciones de pérdida estructuradas explícitamente.

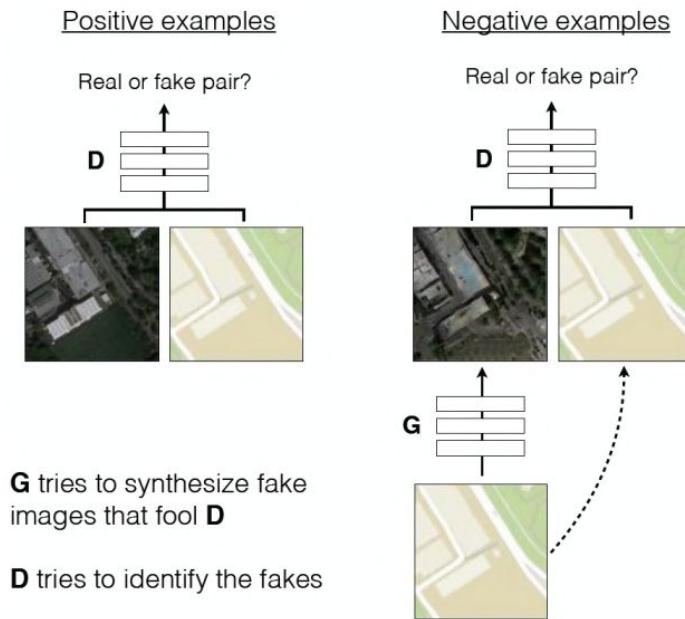


Figure 2 in the original paper.

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. "Image-to-image translation with conditional adversarial networks". arXiv preprint arXiv:1611.07004. (2016).

Aplicación: Traducción de texto a imagen

Motivación:

- Dada una descripción de texto, generar imágenes estrechamente asociadas.
- Utiliza una GAN condicional con el generador y el discriminador condicionados a la incrustación de texto "denso".

this small bird has a pink breast and crown, and black primaries and secondaries.



this magnificent fellow is almost all black with a red crest, and white cheek patch.



the flower has petals that are bright pinkish purple with white stigma



this white and yellow flower have thin white petals and a round yellow stamen



Figure 1 in the original paper.

Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. "Generative adversarial text to image synthesis". ICML (2016).

Aplicación: Traducción de texto a imagen

Text-to-Image Synthesis

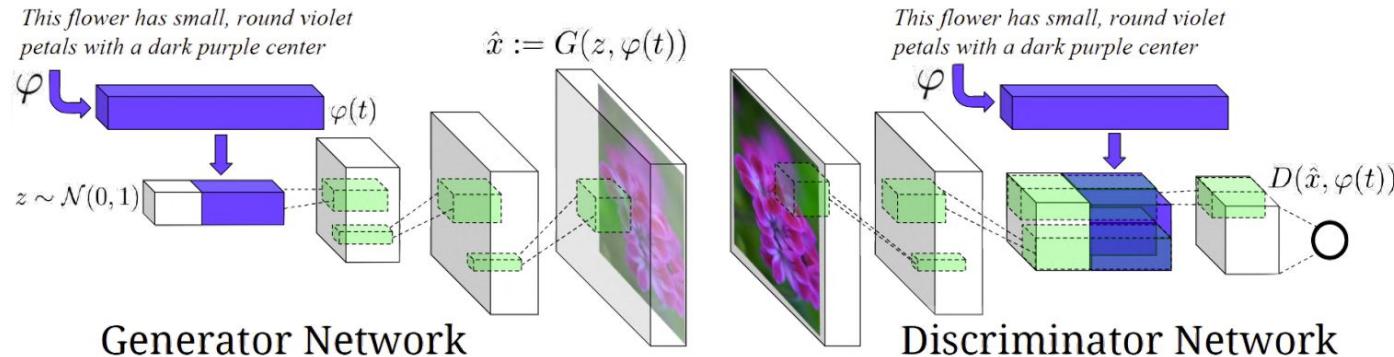


Figure 2 in the original paper.

Positive Example:
Real Image, Right Text

Negative Examples:
Real Image, Wrong Text
Fake Image, Right Text

Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. "Generative adversarial text to image synthesis". ICML (2016).

Aplicación: Envejecimiento facial con GANs condicionadas

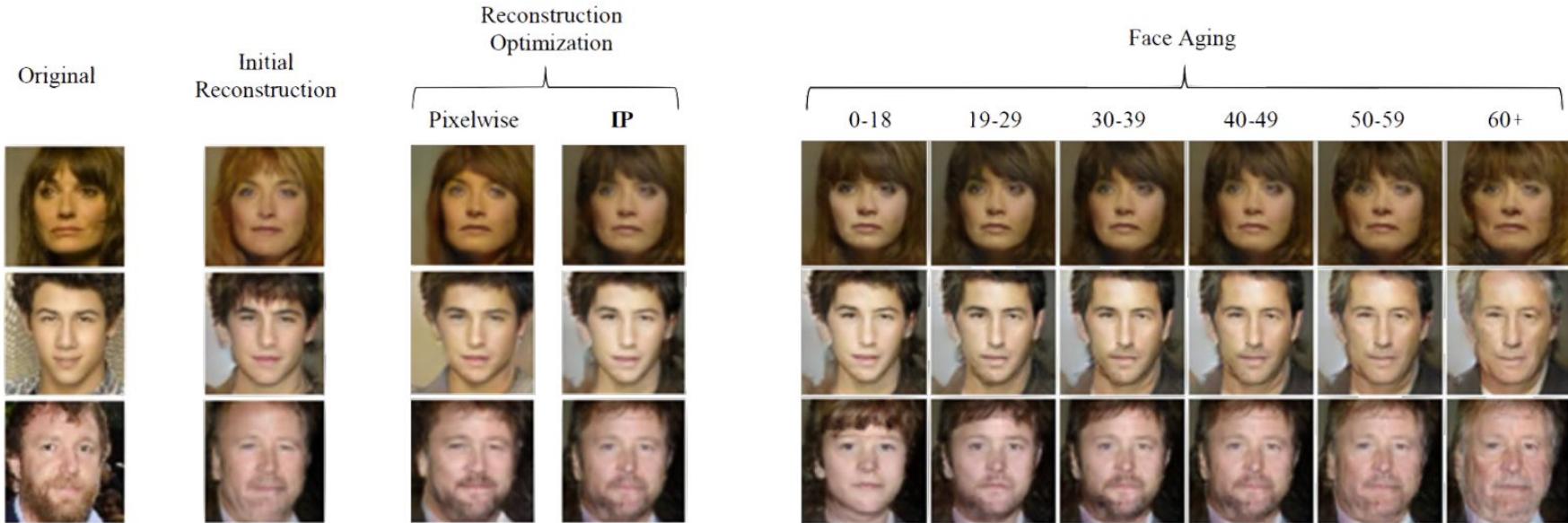
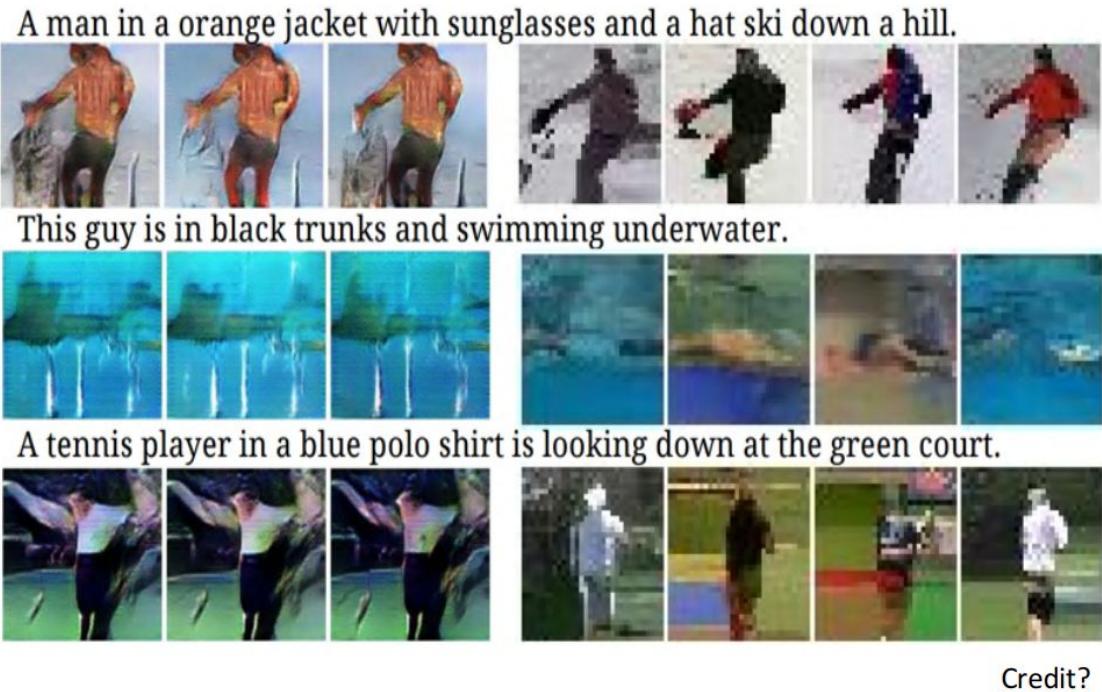


Figure 3 in the original paper.

Antipov, G., Baccouche, M., & Dugelay, J. L. (2017). "Face Aging With Conditional Generative Adversarial Networks". arXiv preprint arXiv:1702.01983.

GANs condicionada en modo-colapso

- Escenario observado cuando la GANs condicionada empieza a ignorar la condición (c) o el espacio latente (z).
- Esto limita la diversidad de imágenes generadas.



Credit?

Mirza, Mehdi, and Simon Osindero. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784 (2014).

Resumen

- Las GANs son modelos generativos que se implementan mediante dos módulos de red neuronal estocástica: Generador y Discriminador.
- El generador intenta generar muestras a partir de ruido aleatorio como entrada.
- El discriminador trata de distinguir las muestras del generador y las muestras de la distribución de datos reales.
- Ambas redes son entrenadas de manera adversa (en conjunto) para engañar al otro componente. En este proceso, ambos modelos mejoran en sus respectivas tareas.

¿Por qué usar GANs para generación?

- Se puede entrenar usando retropropagación para redes neuronales basado en las funciones del Generador/Discriminador.
- Se pueden generar imágenes más nítidas.
- Muestreo más rápido de la distribución del modelo: pase directo único genera una sola muestra.

Lista de lectura

- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. [Generative adversarial nets](#), NIPS (2014).
- Goodfellow, Ian [NIPS 2016 Tutorial: Generative Adversarial Networks](#), NIPS (2016).
- Radford, A., Metz, L. and Chintala, S., [Unsupervised representation learning with deep convolutional generative adversarial networks](#). arXiv preprint arXiv:1511.06434. (2015).
- Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. [Improved techniques for training gans](#). NIPS (2016).
- Chen, X., Duan, Y., Houthooft, R., Schulman, J., Sutskever, I., & Abbeel, P. [InfoGAN: Interpretable Representation Learning by Information Maximization Generative Adversarial Nets](#), NIPS (2016).
- Zhao, Junbo, Michael Mathieu, and Yann LeCun. [Energy-based generative adversarial network](#). arXiv preprint arXiv:1609.03126(2016).
- Mirza, Mehdi, and Simon Osindero. [Conditional generative adversarial nets](#). arXiv preprint arXiv:1411.1784 (2014).
- Liu, Ming-Yu, and Oncel Tuzel. [Coupled generative adversarial networks](#). NIPS (2016).
- Denton, E.L., Chintala, S. and Fergus, R., 2015. [Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks](#). NIPS (2015)
- Dumoulin, V., Belghazi, I., Poole, B., Lamb, A., Arjovsky, M., Mastropietro, O., & Courville, A. [Adversarially learned inference](#). arXiv preprint arXiv:1606.00704 (2016).

Aplicaciones:

- Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. [Image-to-image translation with conditional adversarial networks](#). arXiv preprint arXiv:1611.07004. (2016).
- Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. [Generative adversarial text to image synthesis](#). JMLR (2016).
- Antipov, G., Baccouche, M., & Dugelay, J. L. (2017). [Face Aging With Conditional Generative Adversarial Networks](#). arXiv preprint arXiv:1702.01983.

Bonus Track

Real Images



Fake Images



<https://github.com/ichaparroc/Talk-GANs>

GRACIAS