

Redes generativas adversarias (RGA)

Verónica E. Arriola-Rios

Facultad de Ciencias, UNAM

29 de mayo de 2023



Definición

- 1 Definición
- 2 Arquitectura
- 3 Entrenamiento
- 4 Ejemplo

RG

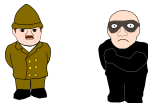
Definición (RGA)

Las *redes generativas adversarias (RGA)*, más populares por su nombre en inglés *Generative Adversarial Networks (GAN)* son una arquitectura compuesta por dos tipos de redes:

Generadora Cualquier red capaz de generar nuevos ejemplares de una clase.

Discriminadora Una red capaz de evaluar si un ejemplar pertenece al conjunto de entradas originales X o si se trata de una imitación.

La red generadora trata de sintetizar ejemplares de la distribución $P(X)$ y la red discriminadora busca distinguir a los ejemplares sintéticos de los reales.



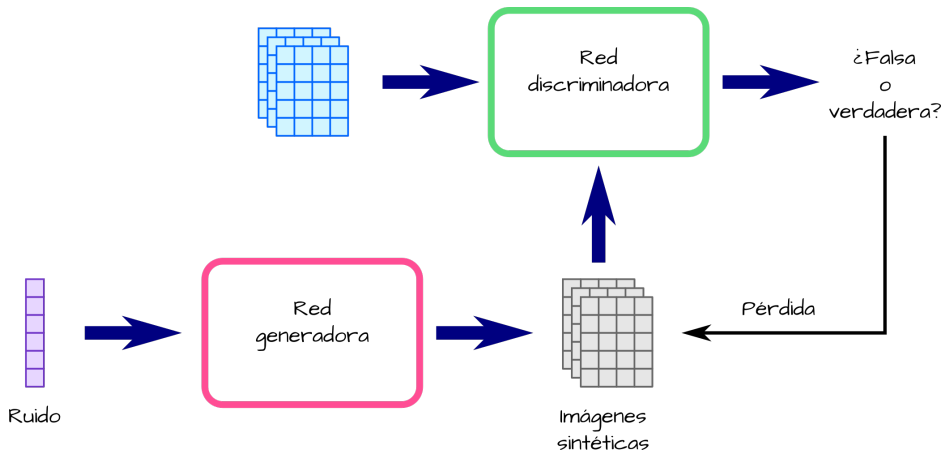
Características

- Se dice que las VAE ya producen mejores resultados, sin embargo la idea de tener dos redes compitiendo (*entrenamiento adversarial*) ha destacado por su ingenio y efectividad.
- Es un ejemplo de modelo *actor-crítico*.

Arquitectura

- 1 Definición
- 2 **Arquitectura**
- 3 Entrenamiento
- 4 Ejemplo

GAN



Entrenamiento

- 1 Definición
- 2 Arquitectura
- 3 Entrenamiento**
- 4 Ejemplo

Entrenamiento

- Ambas redes deben ser entrenadas paulatinamente en forma alternada:
 - 1 Se entrena un poco a la red discriminadora para que distinga a las imágenes X de imágenes con ruido, utilizando **entropía cruzada**.
 - 2 Ahora se entrena a la generadora para que logre engañar a la discriminadora.
- Es importante que ninguna de las dos se vuelva demasiado experta, o la otra no tendrá oportunidad de aprender.
- Al alcanzar el **equilibrio** final, la red generadora deberá haber capturado $P(X)$ y la discriminadora acertará el 50 % de las veces.
- **Debilidad:** La forma más sencilla de engañar a la generadora es aprender a copiar alguno de los ejemplares de entrenamiento, por lo que es difícil obligar a la red generativa a ser creativa.

Ejemplo

- 1 Definición
- 2 Arquitectura
- 3 Entrenamiento
- 4 Ejemplo

Dígitos

Para trabajar con imágenes, como los dígitos de MNIST, las redes son: (Silva 2018)

Discriminadora:



- Red convolucional para clasificación con saltos (*stride*), sin capas de unión (*pooling*) ni completamente conectadas, con normalización por lotes.

- Utiliza *leaky ReLUs* como funciones de activación.

Generativa: Red deconvolucional, esta incrementa el tamaño de las imágenes interpolando los valores intermedios. La última capa usa ReLU.

- Funciona a partir de una imagen pequeña inicial que consiste en ruido normal.
- La deconvolución o convolución transpuesta genera un cuadrado del tamaño del salto (*stride*) en la imagen de salida, por cada pixel en la imagen de entrada.

Referencias I

-  Nicholson, Chris V. (s.f.). *A Beginner's Guide to Generative AI*. URL: <https://wiki.pathmind.com/generative-adversarial-network-gan>.
-  Silva, Thalles (ene. de 2018). *An intuitive introduction to Generative Adversarial Networks (GANs)*. URL: <https://www.freecodecamp.org/news/an-intuitive-introduction-to-generative-adversarial-networks-gans-7a2264a81394>.

Licencia

Creative Commons
Atribución-No Comercial-Compartir Igual

