

# Redes GAN

Redes Convolucionales

Capa SoftMax

GAN: Generative Adversarial Networks

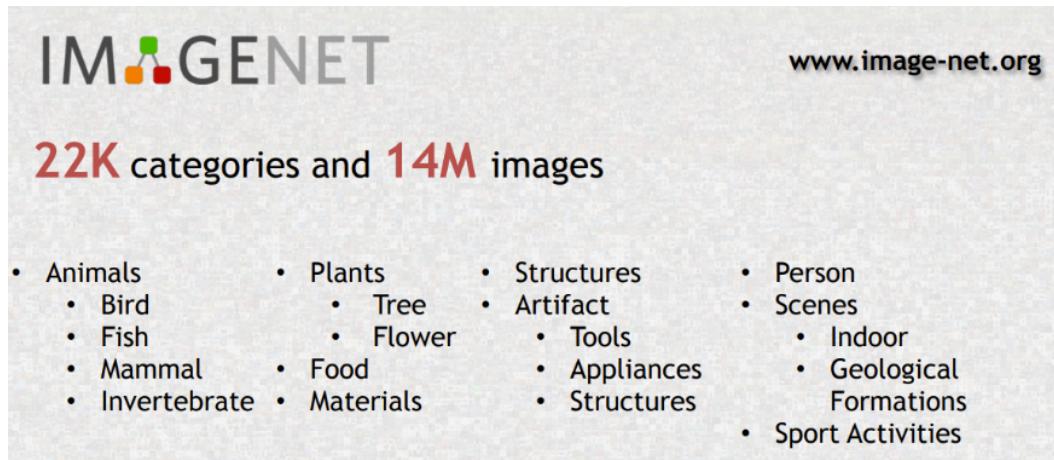
Características

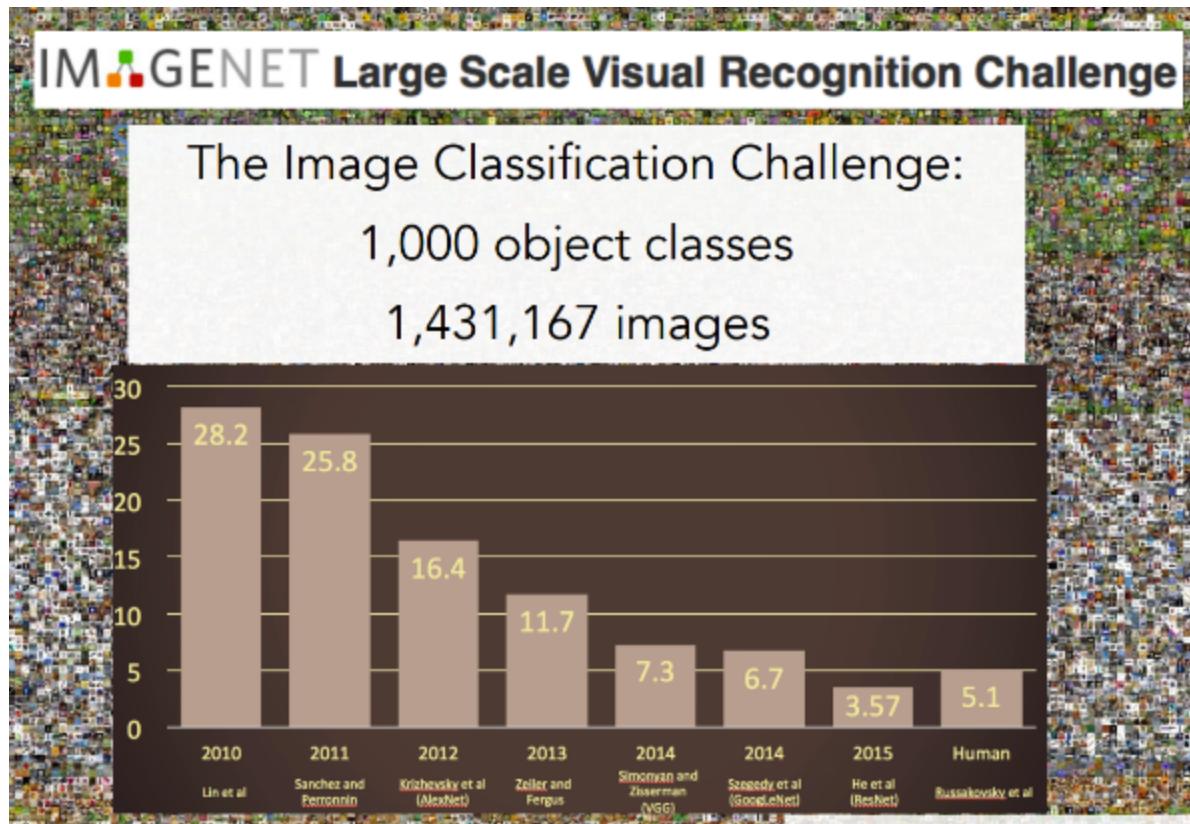
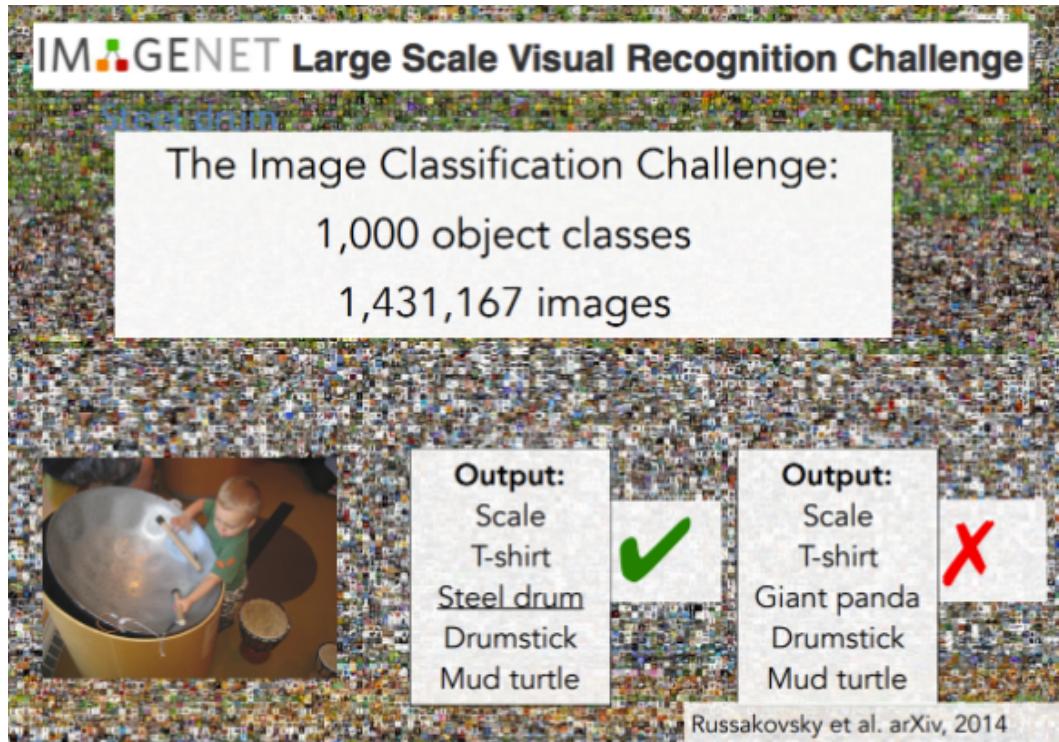
Que es lo que hacen? → ¡Generan nuevos datos!

Funcionamiento

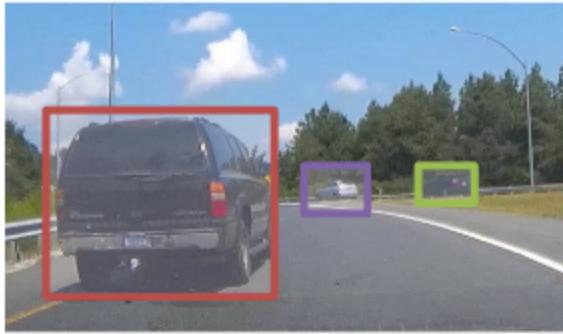
ProGAN

## Redes Convolucionales

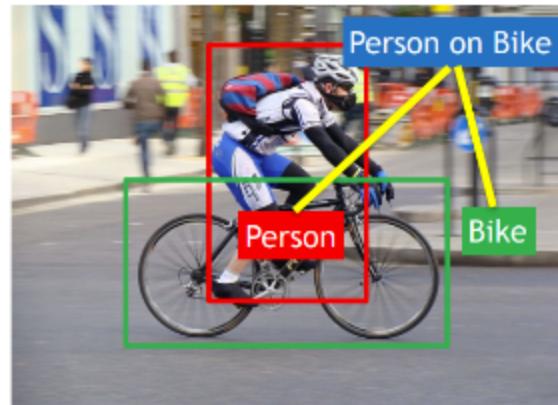




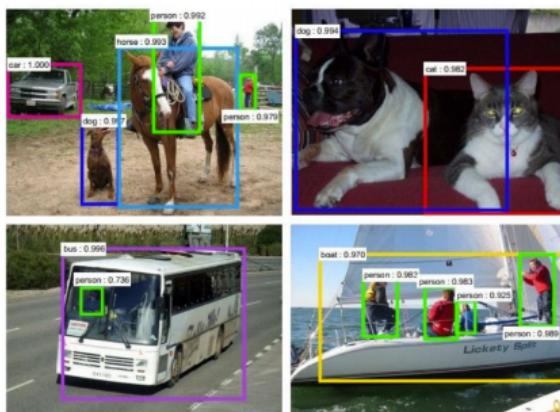
En el 2015 una red neuronal supero el error promedio de los humanos en el reconocimiento de imagenes



This image is licensed under CC BY-NC-SA 2.0; changes made



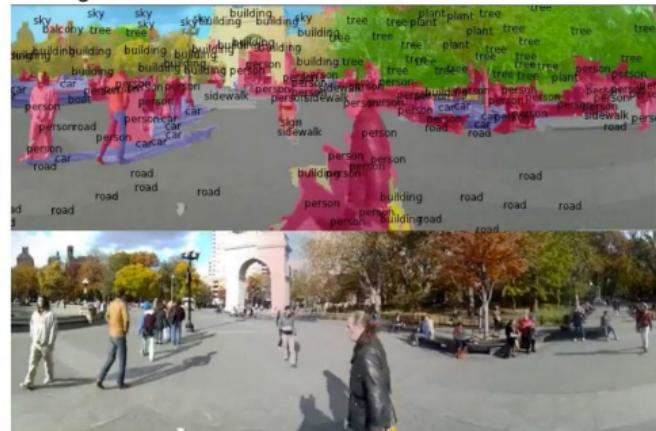
#### Detection



Figures copyright Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, 2015. Reproduced with permission.

**[Faster R-CNN: Ren, He, Girshick, Sun 2015]**

#### Segmentation

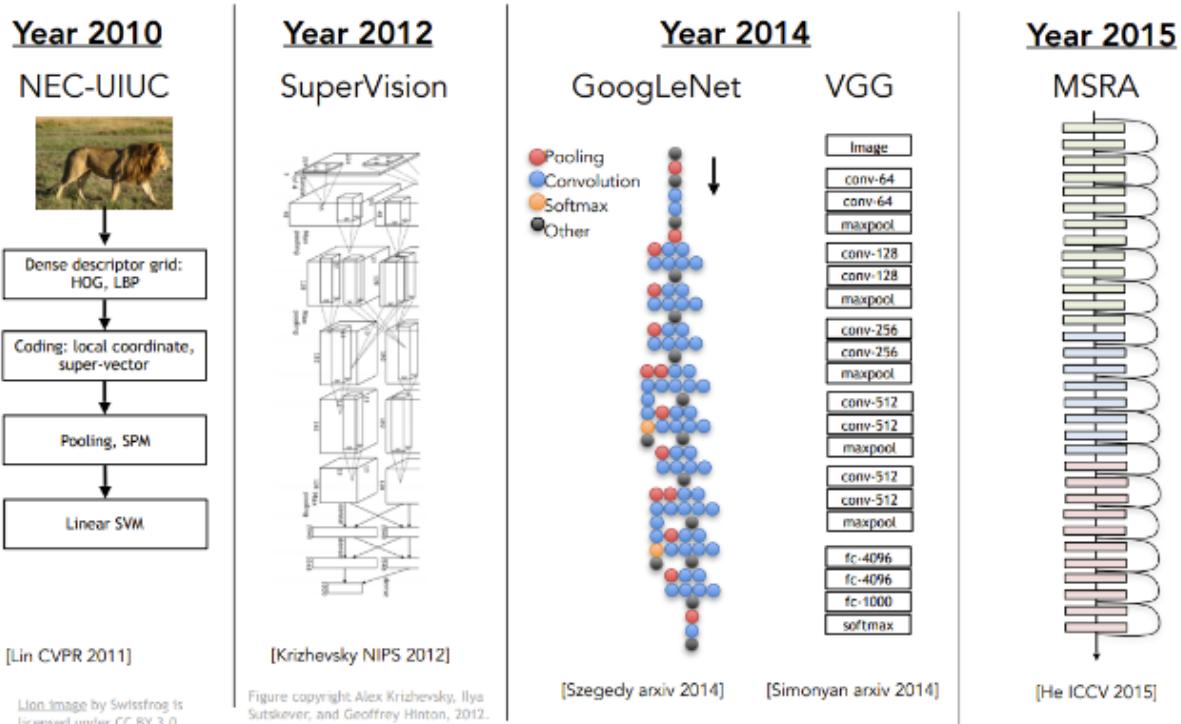


Figures copyright Clement Farabet, 2012.  
Reproduced with permission.

**[Farabet et al., 2012]**

*Fig. 5. Conv nets applications from: Li, F., Johnson, J., Yeung, S. "Stanford CS231n course, Lecture 5", 18 Apr 2017*

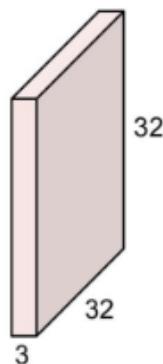
# IMAGENET Large Scale Visual Recognition Challenge



## Repaso de redes convolucionales

### Convolution Layer

32x32x3 image



5x5x3 filter

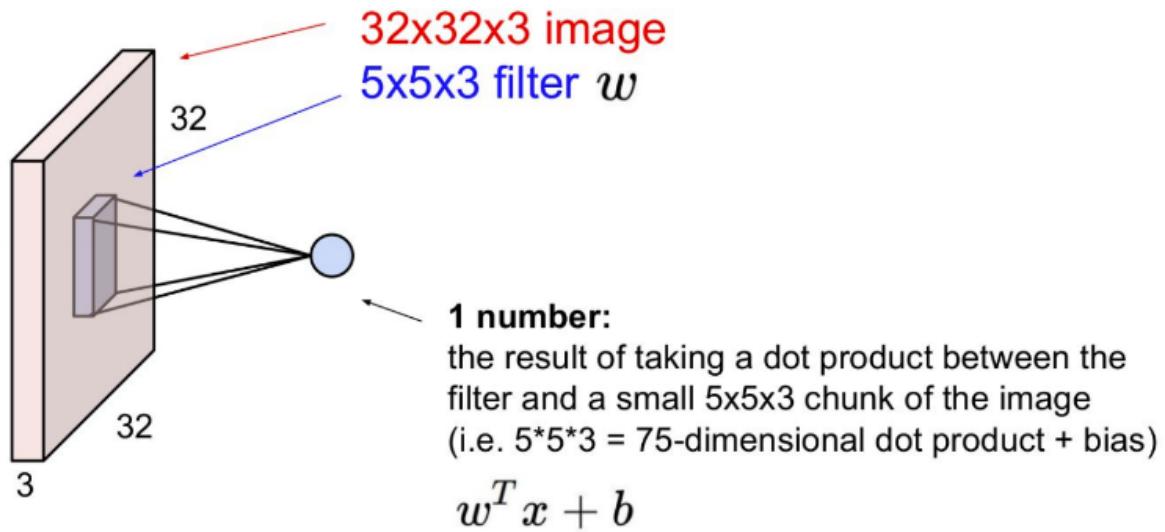


**Convolve** the filter with the image i.e. “slide over the image spatially, computing dot products”

Fig. 6. Convolutional Neural Networks from: Li, F., Johnson, J., Yeung, S. "Stanford CS231n course, Lecture 5". 18 Apr 2017

Imagen de 32x32, 3 canales de profundidad (colores)

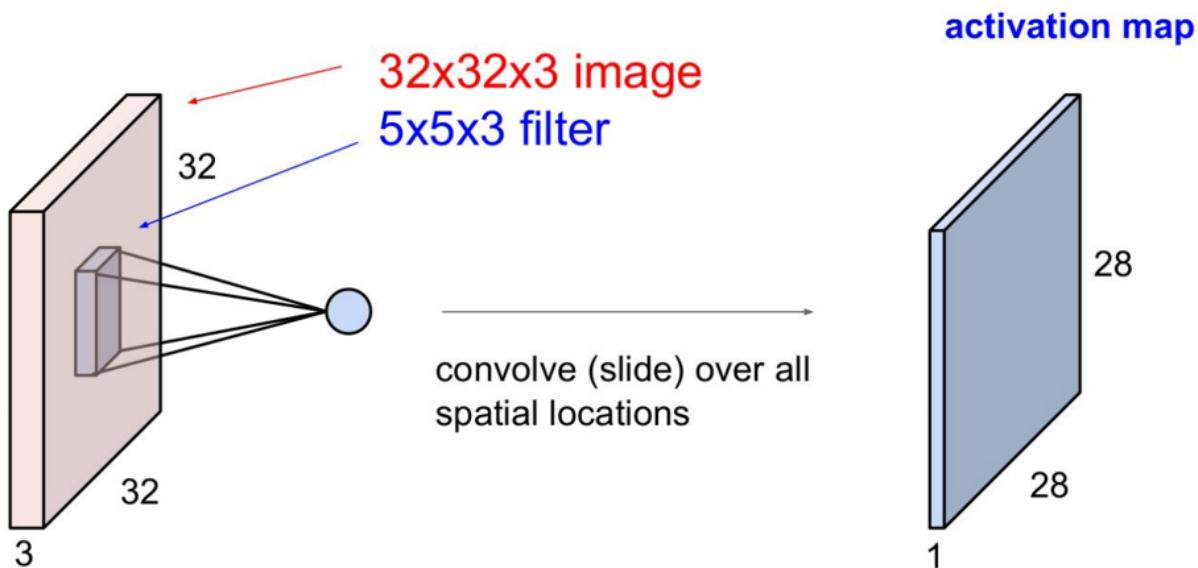
Hay un filtro de 5x5



w: matriz de peso

x: entrada

b: bias



Deja una salida de 28x28

## Convolution Layer

consider a second, green filter

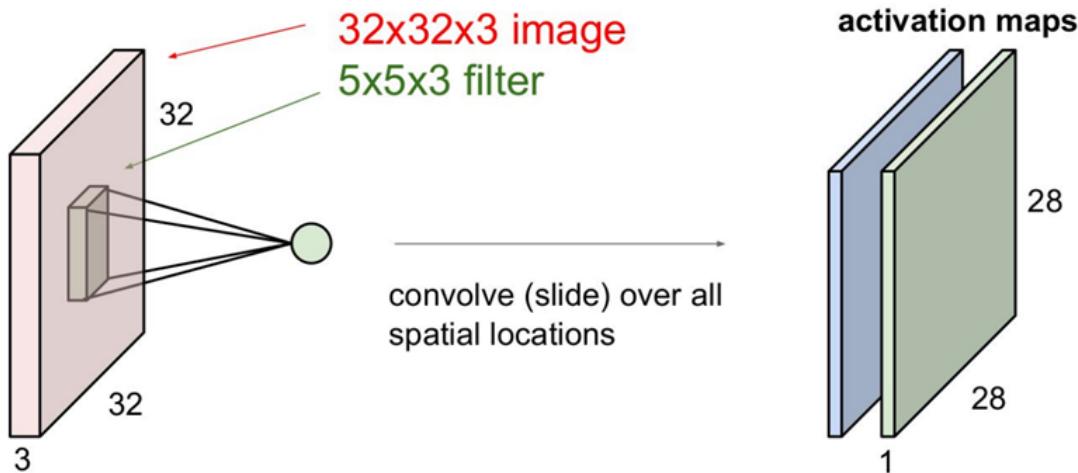
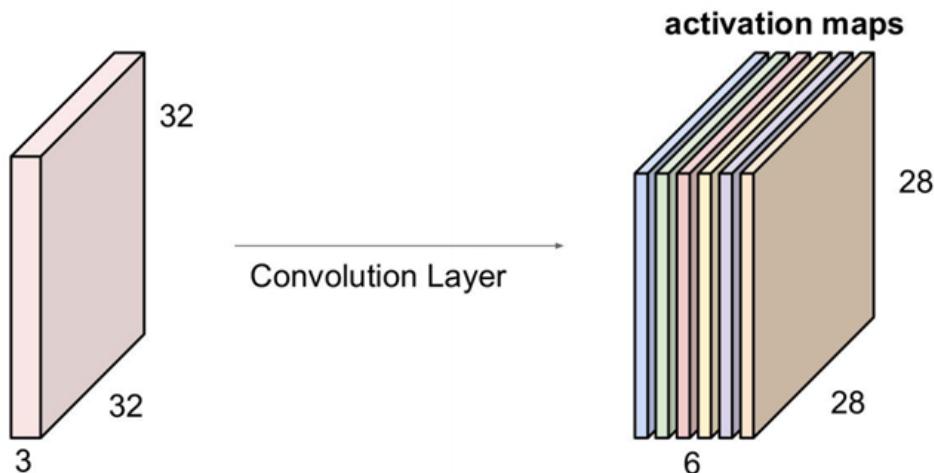


Fig. 9. Convolutional Neural Networks from: Li, F., Johnson, J., Yeung, S. "Stanford CS231n course, Lecture 5", 18 Apr 2017

Como tenemos varios filtros pueden quedar varias imágenes

For example, if we had 6 5x5 filters, we'll get 6 separate activation maps:



We stack these up to get a “new image” of size 28x28x6!

Fig. 10. Convolutional Neural Networks from: Li, F., Johnson, J., Yeung, S. "Stanford CS231n course, Lecture 5", 18 Apr 2017

Varias capas de redes convolucionales

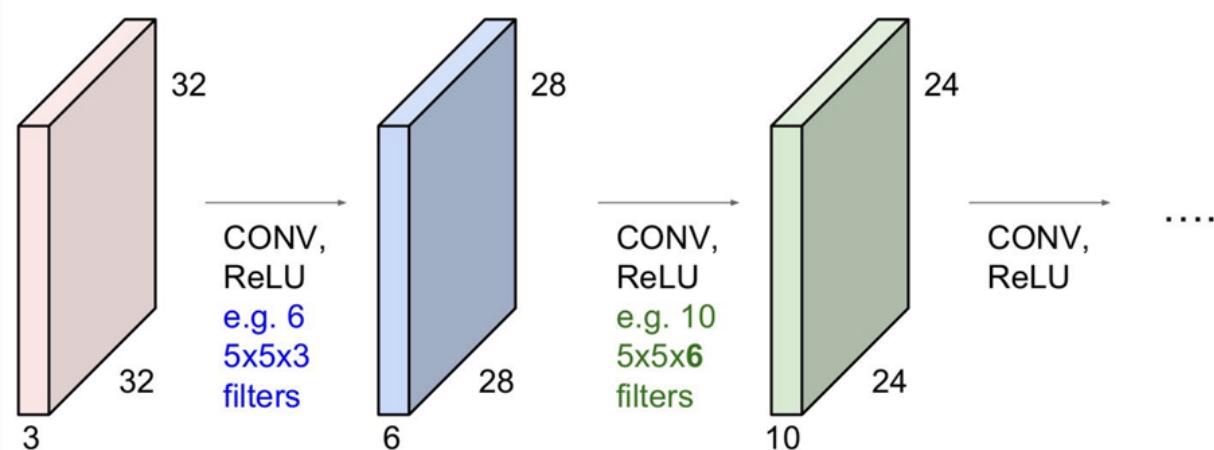
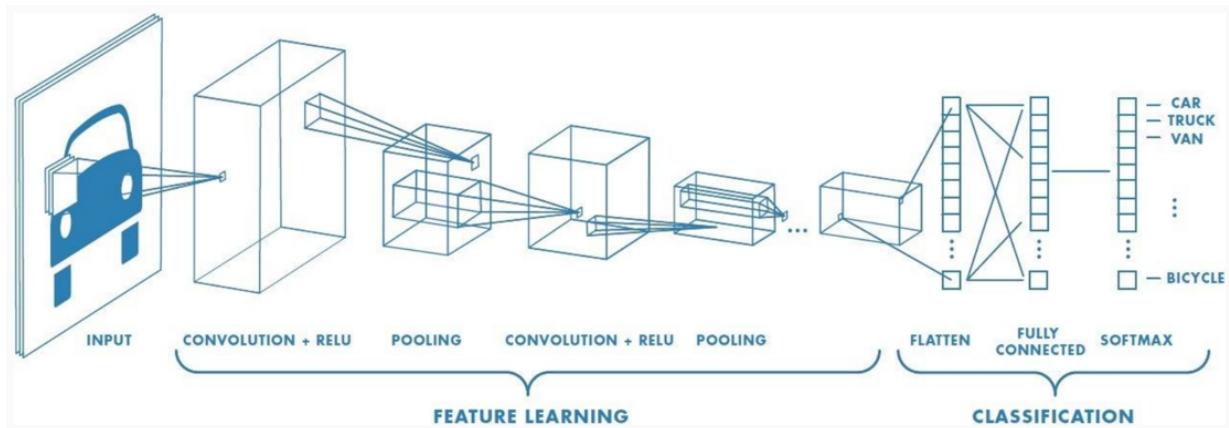


Fig. 11. Convolutional Neural Networks from: Li, F., Johnson, J., Yeung, S. "Stanford CS231n course, Lecture 5", 18 Apr 2017

Y en definitiva, queda algo como lo siguiente



Un pedacito toma una porcion de la imagen

Tenemos una img de salida que es leida por una red onvolucional

Varias capas de convolucional, relu, pooling

Hasta que al final tenemos la red de clasificacion que es un perceptron multicapa con la funcion de activacion softmax

## Capa SoftMax

Esta pensada para multiples salidas, se pueden activar varias neuronas a la vez.

**Softmax** lleva la idea de la regresión logística al plano de las clases múltiples.

Asigna probabilidades decimales a cada clase en un caso de clases múltiples. Esas probabilidades decimales deben sumar 1.0. Esta restricción adicional permite que el entrenamiento converja más rápido.

Da un puntaje de que tan segura esta la red de que es un determinado objeto.

Es probabilistico. Si se activan x neuronas con diferentes probabilidades, esas x probabilidades deberian sumar 1.

Clase	Probabilidades
manzana	0.001
oso	0.04
caramelos	0.008
perro	0.95
huevo	0.001

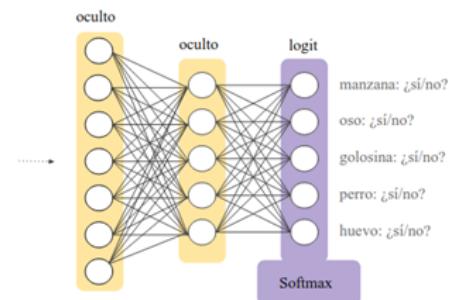


Figura 2. Una capa de softmax en una red neuronal

La neurona ganadora es la de perro.

## GAN: Generative Adversarial Networks

<https://thispersondoesnotexist.com/>

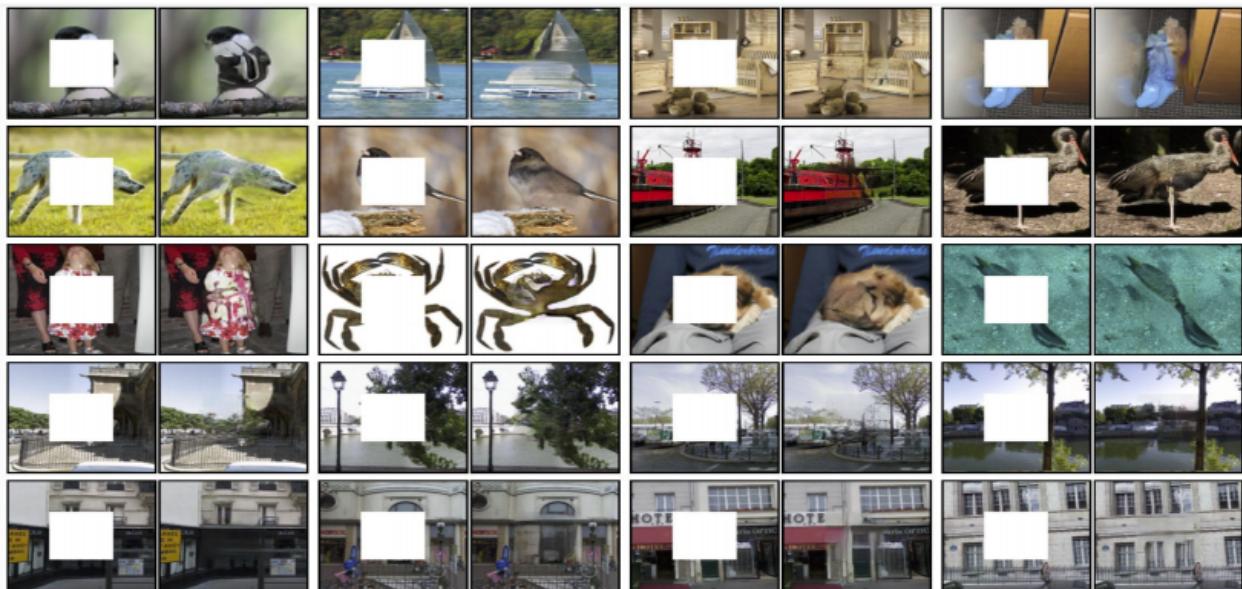
Genera caras aleatorias, las dibuja

## Características

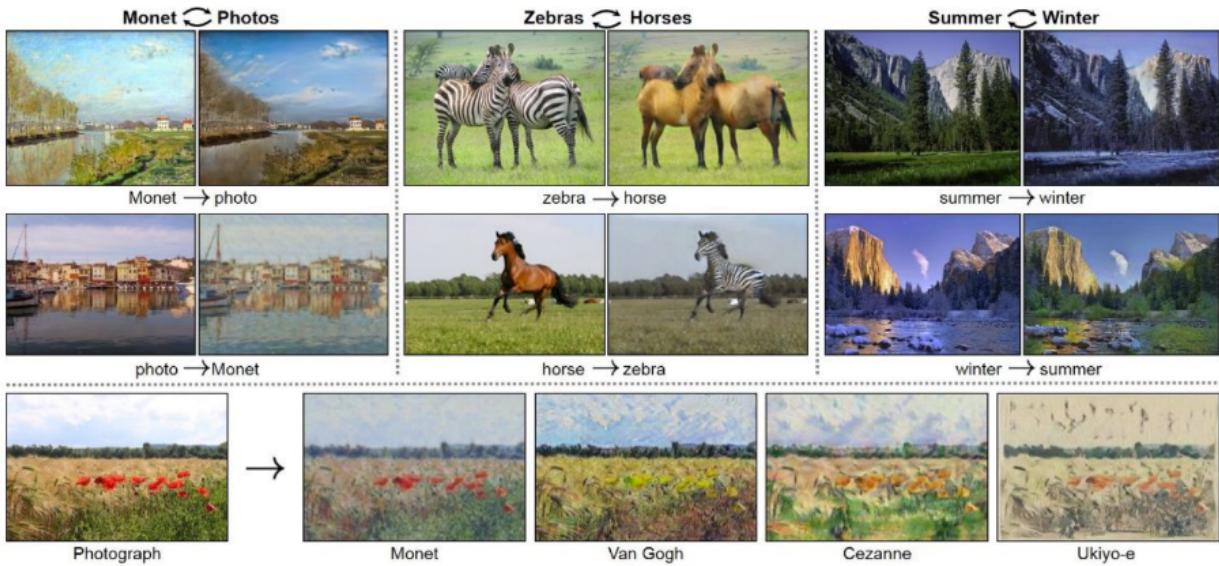
- Son redes NO-Supervisadas  
Hay que entrenarlas, pero no hay que etiquetarlas
- No arman *clusters*
- No clasifican datos  
No indican si es hombre, mujer, pelado, etc
- No resuelven problemas de regresión

## Que es lo que hacen? → ¡Generan nuevos datos!

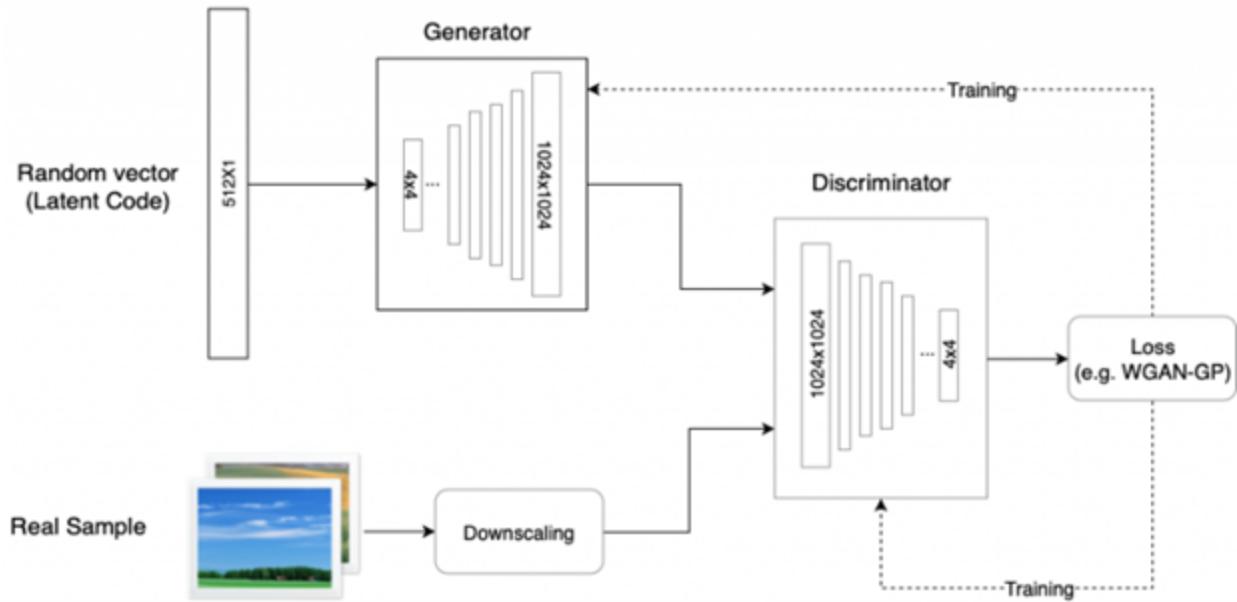
- Completar una imagen a la que le falta un pedazo



- Convertir cualquier foto en un cuadro de un pintor
- O al revés, darle un cuadro y que lo convierta a una img realista
- Convertir caballos a cebras o al revés
- Convertir un paisaje de invierno a verano



## Funcionamiento



Discriminador es la red GAN, la red convolucional

Va a recibir imgs reales de una gran base de datos

Downscaling: se procesan, recortan, se les da una resolucion acorde (1024x1024 en este caso)

Las pasamos por la red conv. y me va a decir si es una foto o no lo es  
(ya fue entrenada esta red)

Generador es una red conv dada vuelta, es una red donde le ingreso ruido por la salida  
y genera una img de ruido

El discriminador me va a decir que img es real y cual fue creada por el generador

Va a detectar a funcion de error, va a hacer una diferencia entre lo que espera y lo que se mando, y va a retroalimentar al generador para que actualice todos sus pesos

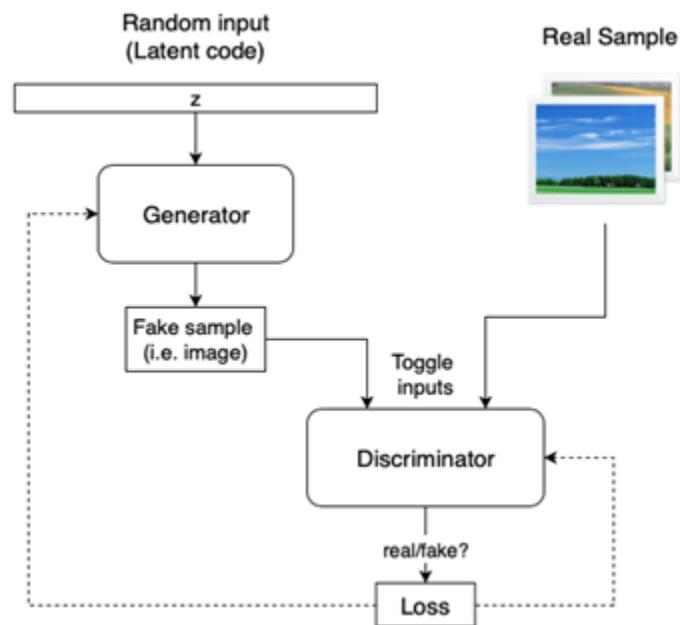
Cuando el generador actualiza todos sus pesos, va a generar otra img

Hay varias capas en el generador y en el discriminador

Generador va de 4x4 a 1024x1024, y el discriminador va de 1024x1024 a 4x4

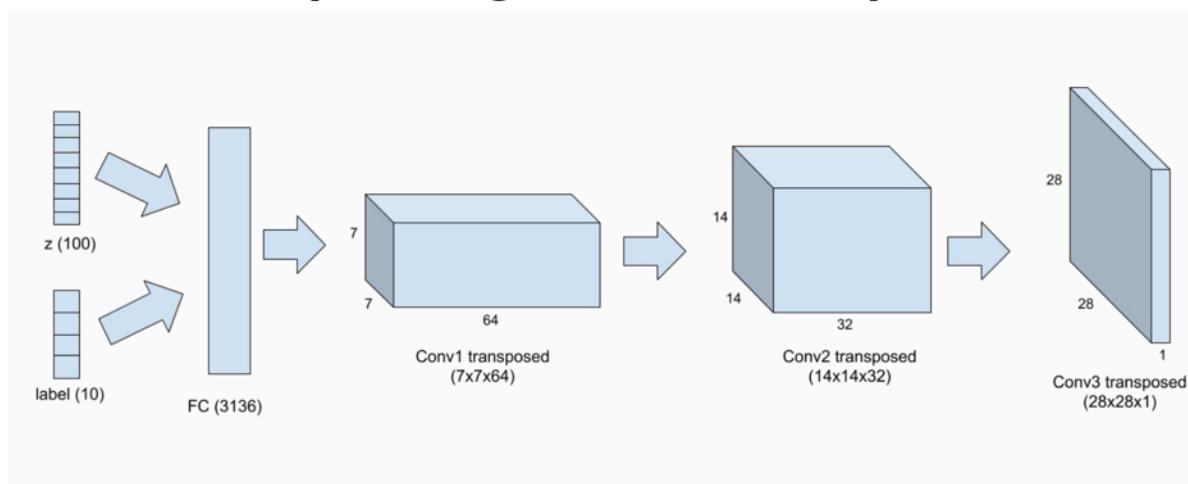
En el primer paso vamos a generar una img de 4x4, y recien cuando el generador logre hacer imgs de 4x4 que engañan al discriminador  $\Rightarrow$  vamos a pasar a imgs de 8x8

El discriminador tambien actualiza sus pesos a medida que se vaya equivocando



Esto es como funciona la red convolucional inversa, la transpuesta

## Generador para dígitos de 28x28 pixeles

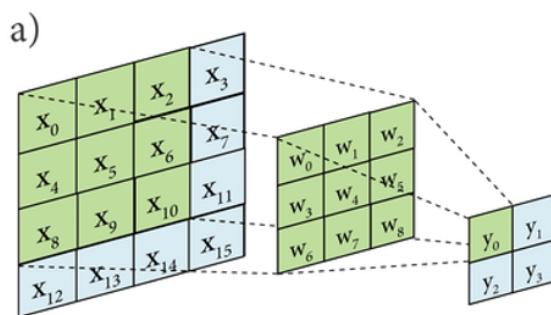


Empezamos con la red Fully Connected al principio que recibe un vector  $z$  de ruido y una etiqueta de lo que se esta generando (vector que se va activando)

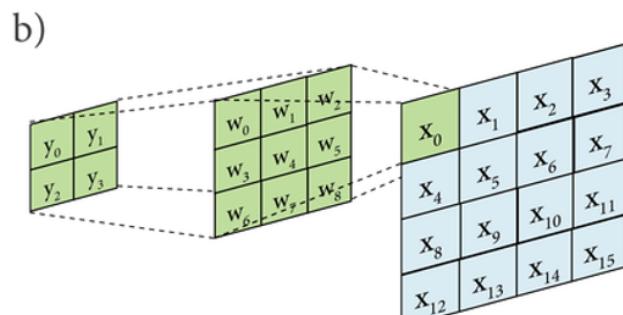
Despues red convolucional transpuesta

Cada vez se van generando imgs mas grandes

## Red Convolucional transpuesta



Convolucional

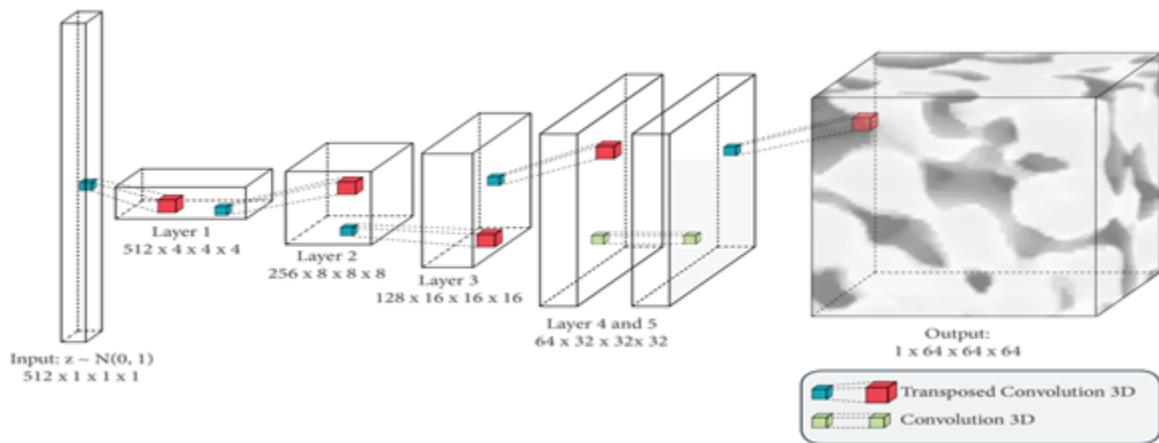


Convolucional Transpuesta

Convolucional: Tomamos una porcion de la img, la multiplicamos por los pesos y sumatorial, y con eso ponemos un valor en la salida

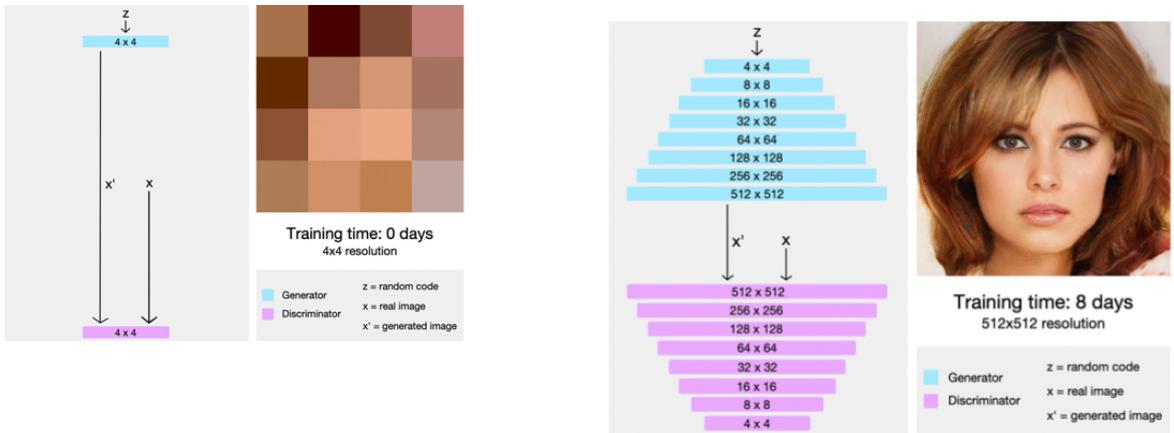
Convolucional Transpuesta: Tomamos la salida, la multiplicamos por los pesos y se obtiene un valor

Capa a capa



## ProGAN

Genera imágenes de alta calidad pero, como en la mayoría de los modelos, su capacidad para controlar características específicas de la imagen generada es muy limitada. En otras palabras, las funciones están enredadas y, por lo tanto, intentar modificar la entrada, aunque sea un poco, generalmente afecta a varias funciones al mismo tiempo. Una buena analogía para eso serían los genes, en los que cambiar un solo gen podría afectar múltiples rasgos



## Synth

Synth.run

 <https://app.synth.run/#/home>

## DALL·E 2

DALL·E 2 is a new AI system that can create realistic images and art from a description in natural language.

 <https://openai.com/dall-e-2/>