

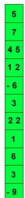
El aprendizaje profundo de la imagen

Jorge Calvo Martín
Profesor Colegio Europeo de Madrid y Universidad Alfonso X

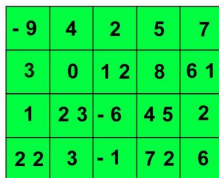


Cómo la máquina ve una imagen

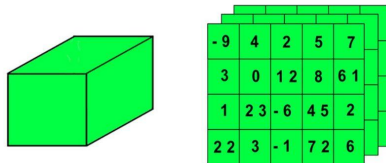
1D TENSOR /
VECTOR



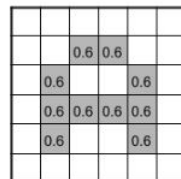
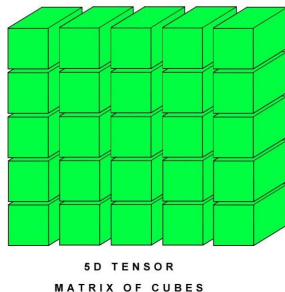
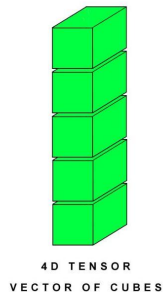
2D TENSOR /
MATRIX



3D TENSOR /
CUBE



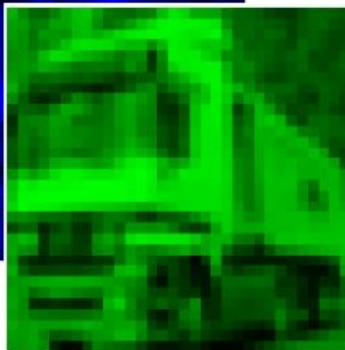
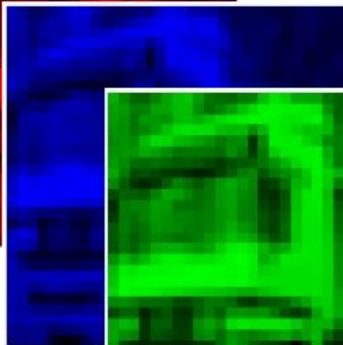
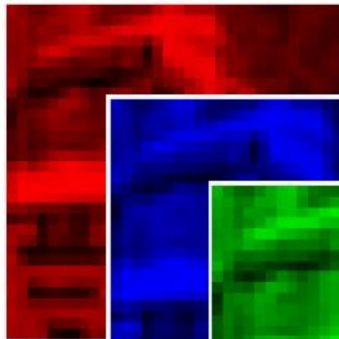
Una imagen...



...es una matriz de pixeles.
El valor de los pixeles va de 0 a 255 pero se normaliza para la red neuronal de 0 a 1



Y el color



		0.2	0.2	
	0.2			0.2
	0.2	0.2	0.2	0.2
	0.2			0.2

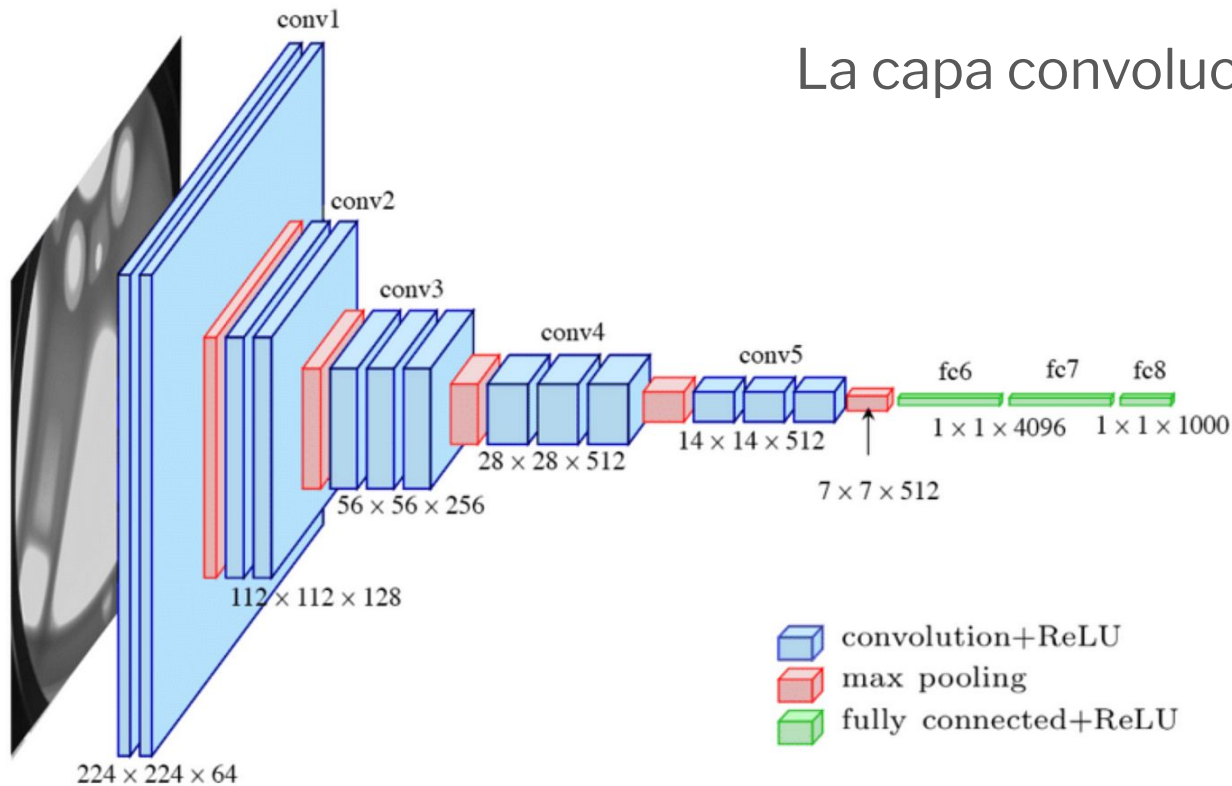
		0.4	0.4	
	0.4			0.4
	0.4	0.4	0.4	0.4
	0.4			0.4

		0.2	0.2	
	0.2			0.2
	0.2	0.2	0.2	0.2
	0.2			0.2

¿Cómo entiende la imagen?



La capa convolucional y las dimensiones



¿Qué hace una Convolución?



Extraer características de una imagen usando las dimensiones y compresión

		0.6	0.6		
	0.6			0.6	
	0.6	0.6	0.6	0.6	
	0.6			0.6	

Imagen de
entrada

1	0	-1
2	0	-2
1	0	-1

kernel



		0.6	0.6		
	0.6			0.6	
	0.6	0.6	0.6	0.6	
	0.6			0.6	

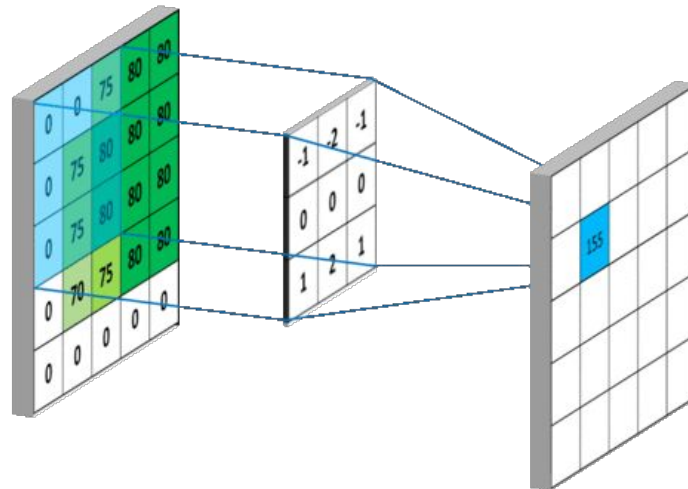
-1,2			



¿Se comprime la imagen?

3 ₀	3 ₁	2 ₂	1	0
0 ₂	0 ₂	1 ₀	3	1
3 ₀	1 ₁	2 ₂	2	3
2	0	0	2	2
2	0	0	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0



¿Mantengo el tamaño? Padding



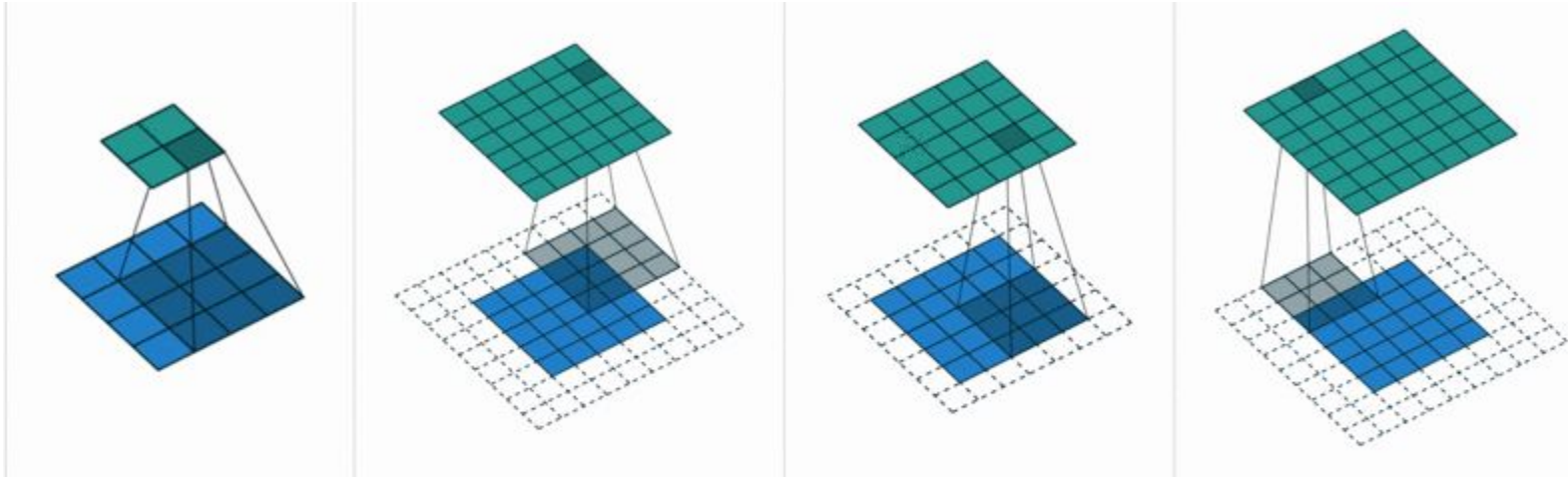
0	0	0	0	0	0	0
0	60	113	56	139	85	0
0	73	121	54	84	128	0
0	131	99	70	129	127	0
0	80	57	115	69	134	0
0	104	126	123	95	130	0
0	0	0	0	0	0	0

Kernel

0	-1	0
-1	5	-1
0	-1	0

114				

¿Mantengo el tamaño? Padding





Vamos a practicar



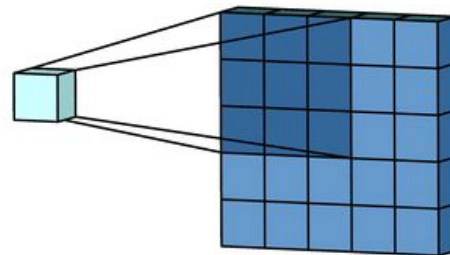
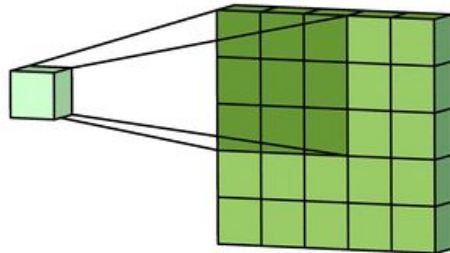
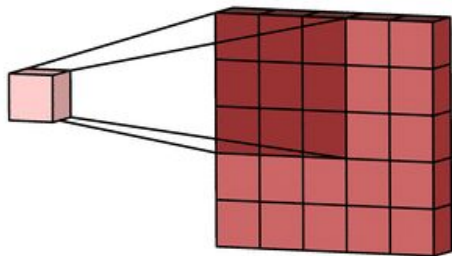
<https://deeplizard.com/resource/pavq7noze2>



<https://setosa.io/ev/image-kernels/>



Y el color



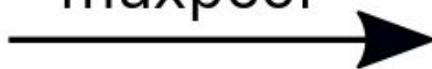


¿Hay otras técnicas?

Input

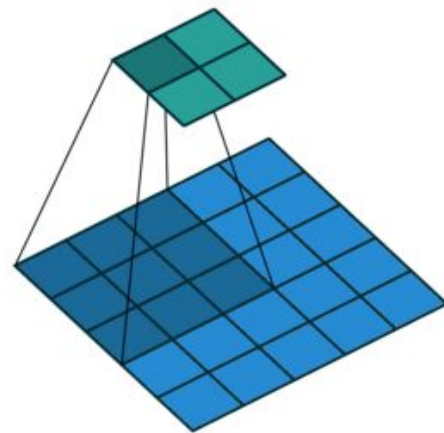
7	3	5	2
8	7	1	6
4	9	3	9
0	8	4	5

maxpool



Output

8	6
9	9





¿Hay otras técnicas?

Strided convolution

2 ³	3 ⁴	7 ⁴	4	6	2	9
6 ¹	6 ⁰	9 ²	8	7	4	3
3 ⁻¹	4 ⁰	8 ³	3	8	9	7
7	8	3	6	6	3	4
4	2	1	8	3	4	6
3	2	4	1	9	8	3
0	1	3	9	2	1	4

7x7

*

3	4	4
1	0	2
-1	0	3

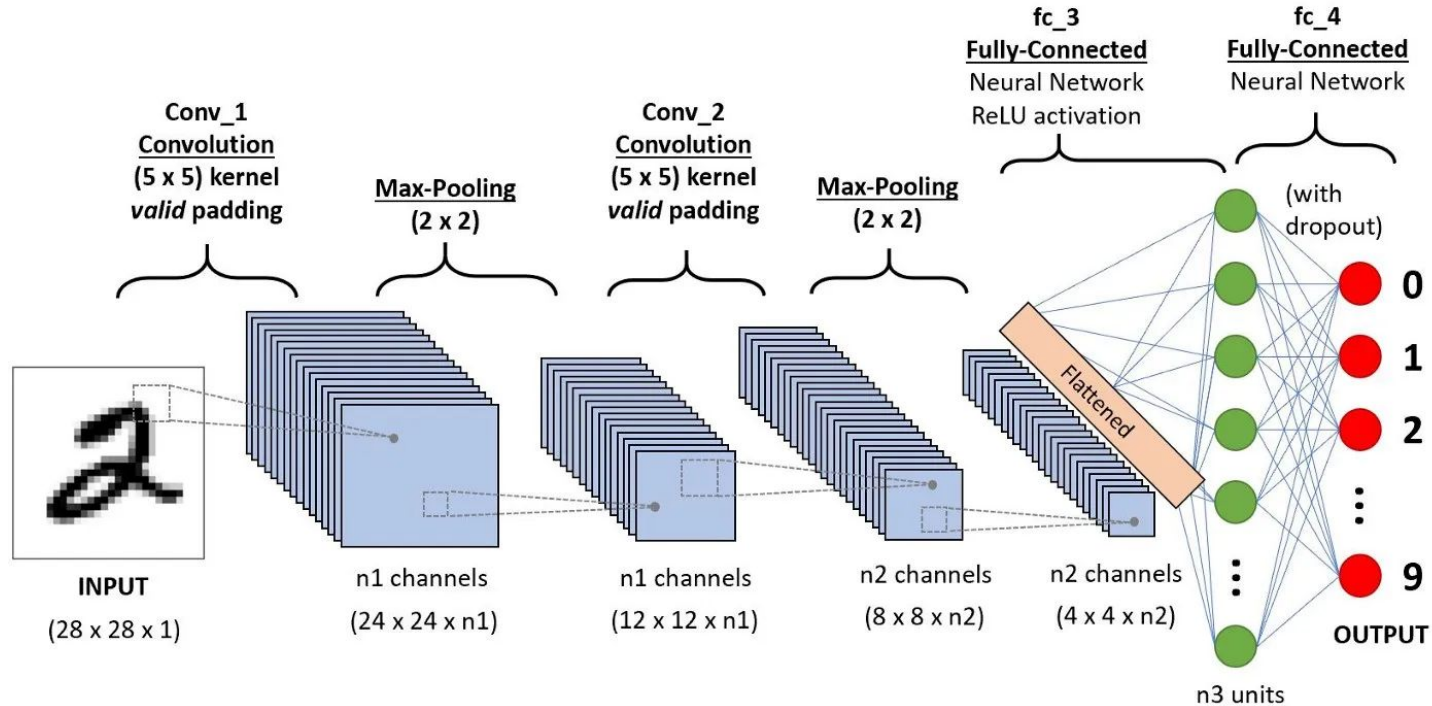
3x3

Stride = 2

=

91		

Aplanar para decidir flatten



Aplanar para decidir flatten



1	1	0
4	2	1
0	2	1

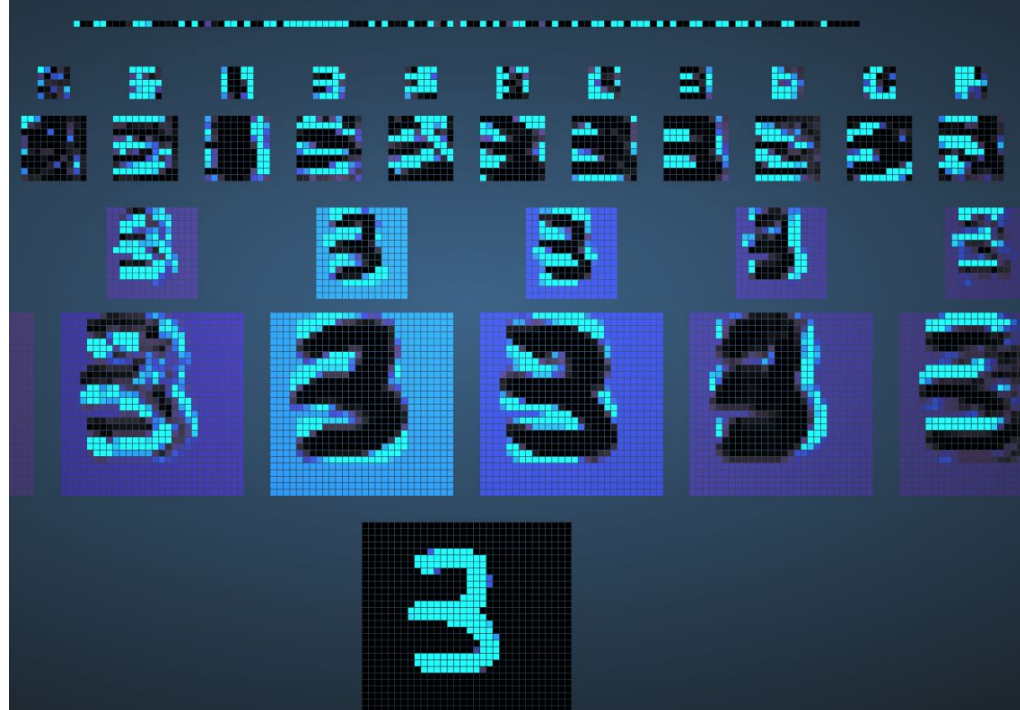
Pooled Feature Map

Flattening



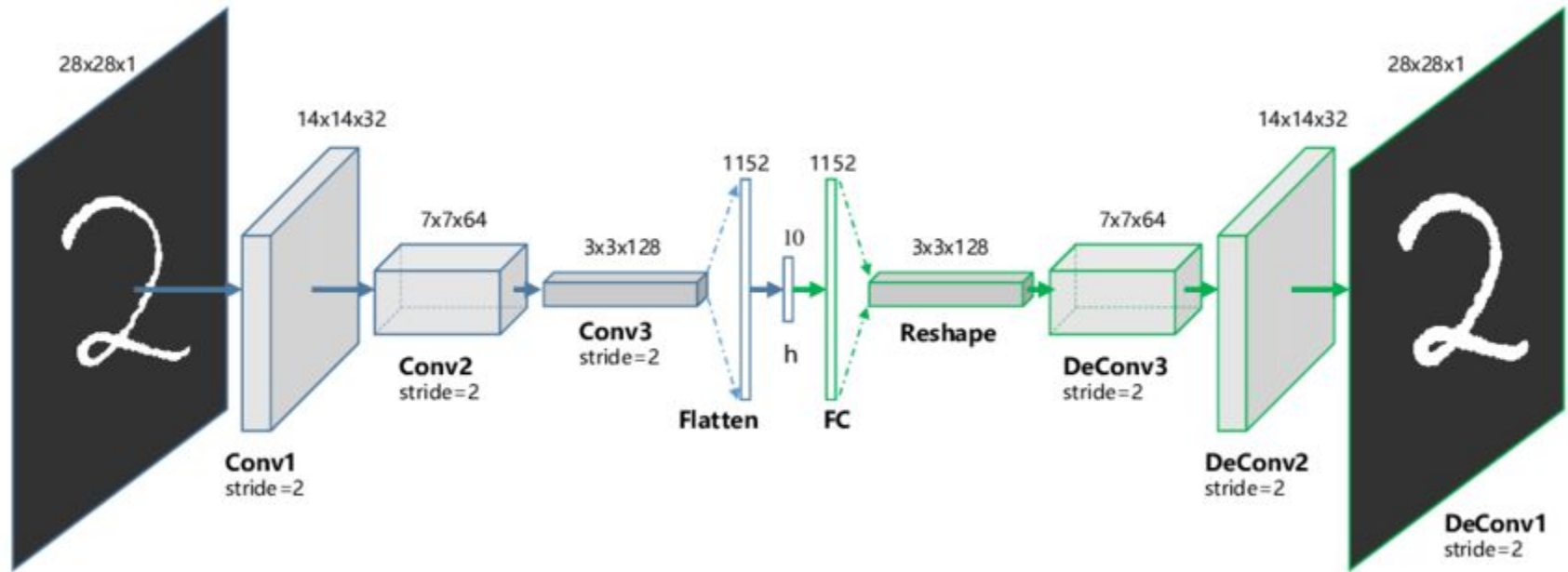
1
1
0
4
2
1
0
2
1

Visualizamos una red neuronal covolucional

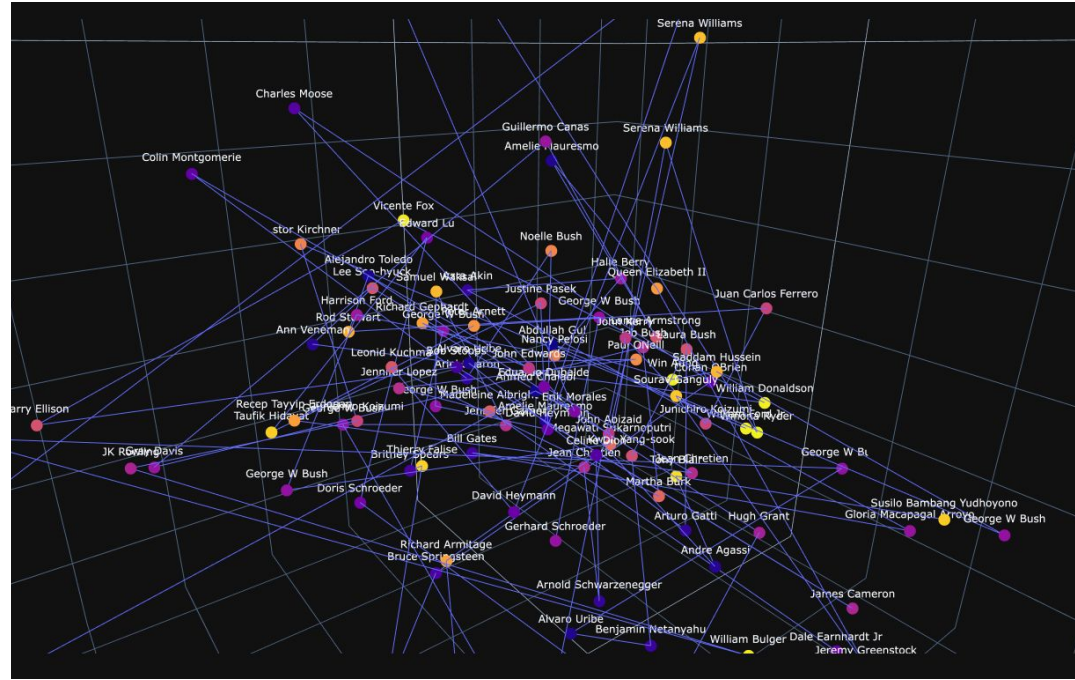


https://adamharley.com/nn_vis/cnn/2d.html

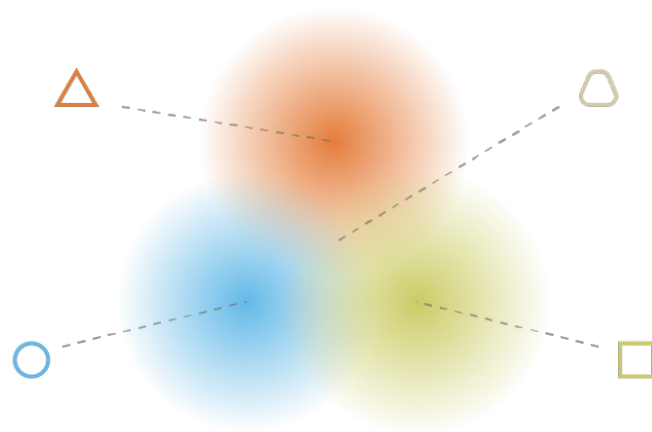
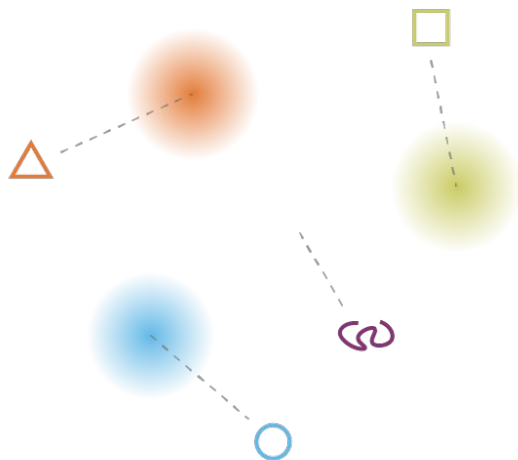
El nacimiento de la IA Generativa en imágenes



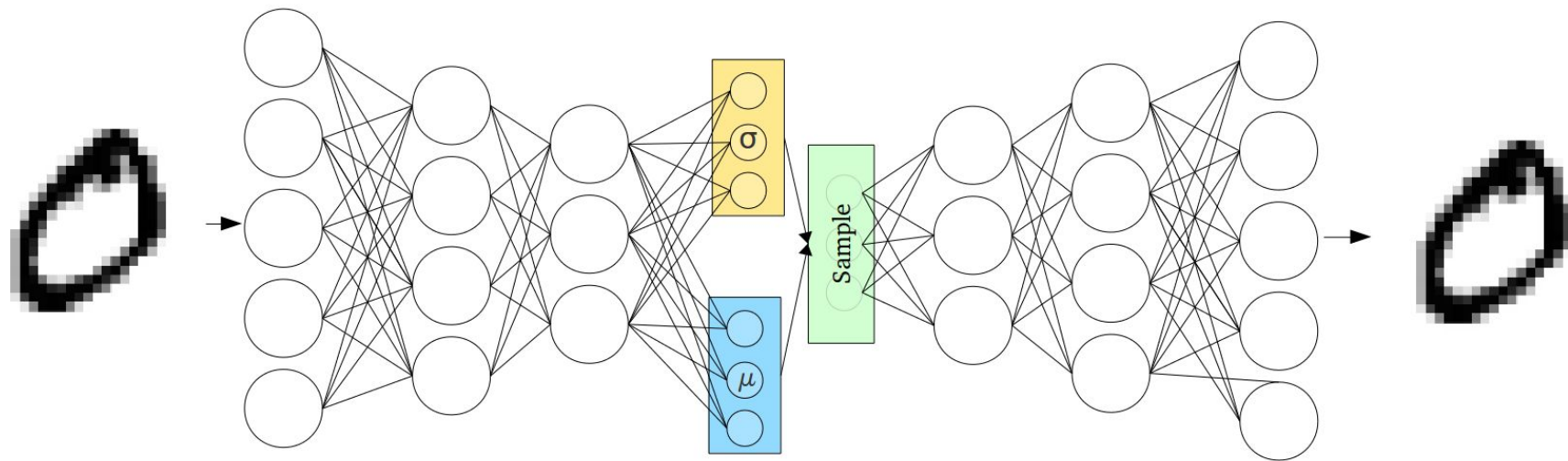
AutoEncoder



De copiar a generar algo nuevo: El espacio latente



De un valor determinista a un valor estocástico



Output μ [0.1, 1.2, 0.2, 0.8,...]

Output σ [0.2, 0.5, 0.8, 1.3,...]

Intermediate X [$X_1 \sim N(0.1, 0.2^2)$, $X_2 \sim N(1.2, 0.5^2)$, $X_3 \sim N(0.2, 0.8^2)$, $X_4 \sim N(0.8, 1.3^2)$, ...]

↓ sample

Sampled vector [0.28, 1.65, 0.92, 1.98,...]



GRACIAS



@jorgemcalvo



jorgecalvomartin



<https://www.europeanvalley.es/noticias>



jmcalvomartin