

REDES NEURONALES CONVOLUCIONALES

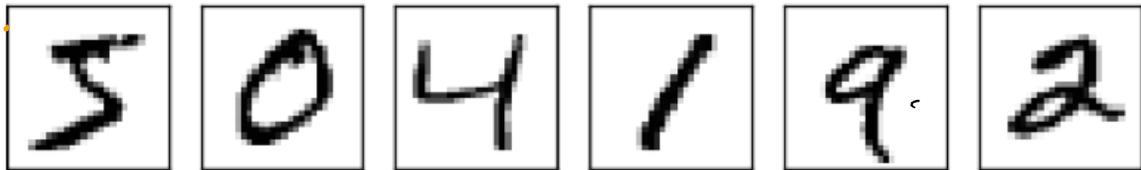
Una red neuronal convolucional es una red de arquitectura feed-forward con aprendizaje SUPERVIZADO. Si bien hoy se usan para una gran diversidad de problemas, surgieron como imitación del sistema visual en mamíferos, y por lo tanto son muy adecuados para el procesamiento de imágenes.

El primer intento se llamó **Neocognitron** y fue introducido por Kunihiko Fukushima en 1980. Pero recién en 1998, Yann LeCun y su equipo introdujeron la versión actual, con back propagation.

Nosotros usaremos un clasificador de imágenes como ejemplo, pero no es necesario que sea así.

Supongamos que queremos reconocer los dígitos 0, 1, 2, ..., 9 usando una red neuronal. Nuestra base de datos será MNIST, que cuenta con 70000 imágenes en tonos de gris de dígitos escritos a mano por voluntarios (NIST).

504192

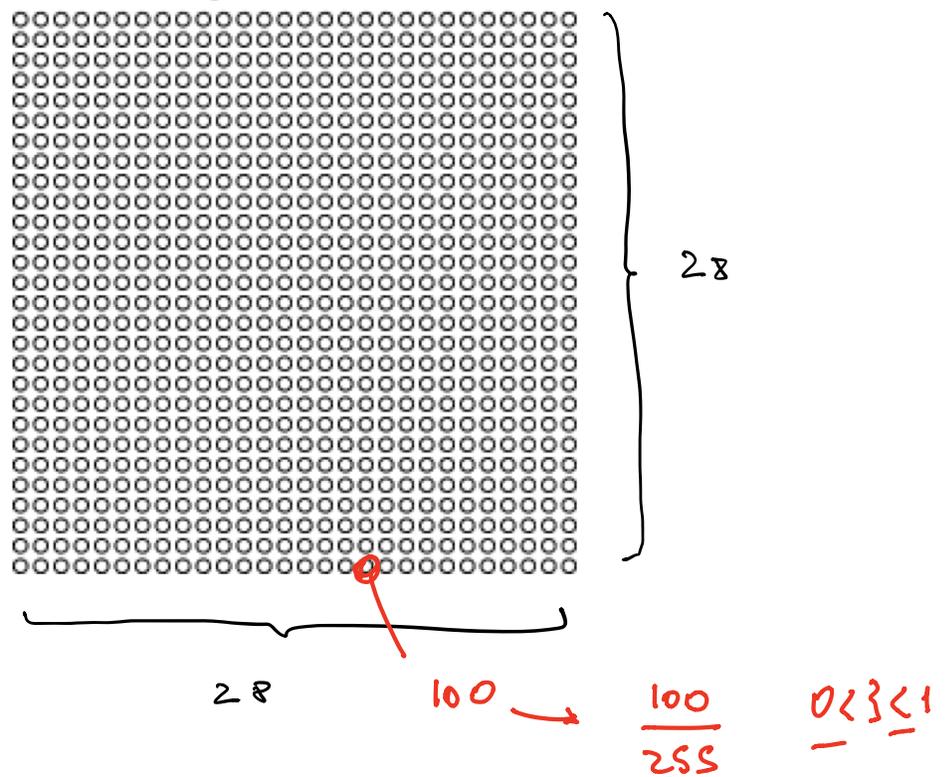


La entrada de la red será una imagen de 28 píxeles de alto y 28 píxeles de ancho, o sea, 784 píxeles, en ~~128~~²⁵⁶ tonos de gris.

Queremos preservar la topología bidimensional de la imagen, o sea, llevar en cuenta la idea de proximidad y posición relativa entre píxeles.

CAPA DE ENTRADA O CAMPO RECEPTIVO

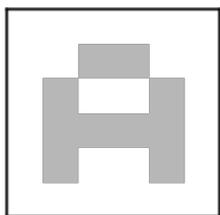
Esta capa de entrada emula el campo receptivo de nuestra retina



Recordemos que el descenso por el gradiente es más eficiente si normalizamos los datos. Así que cada neurona de la capa de entrada o campo receptivo produce tomar valores

$$\frac{g_{in}}{255}$$

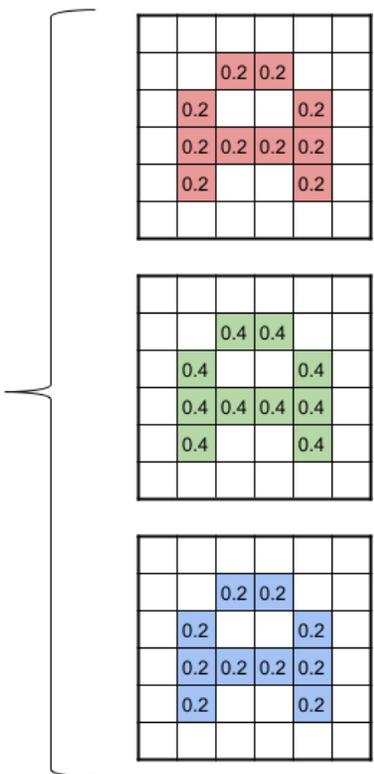
donde $gris = 0, 1, 2, \dots, 255$ es la escala de grises.



Una imagen...

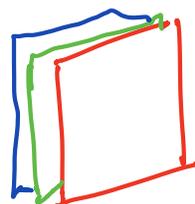
| | | | | | | | |
|--|--|-----|-----|-----|-----|-----|--|
| | | 0 | 0 | 0 | | | |
| | | | 0.6 | 0.6 | | | |
| | | 0.6 | | | | 0.6 | |
| | | 0.6 | 0.6 | 0.6 | 0.6 | | |
| | | 0.6 | | | | 0.6 | |
| | | | | | | | |

...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



Si la imagen es a color, estará compuesta de tres canales: rojo, verde, azul.

RGB



28
x (28 x 28)
451584

En este caso tenemos una entreda tridimensional

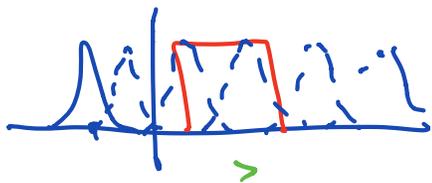
$$28 \times 28 \times 3$$

CAPA CONVOLUCIONAL

Una convolución es un filtro

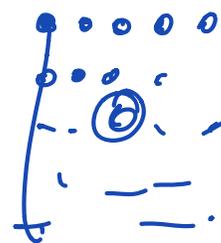
En matemática, una convolución es una operación de dos funciones, f y g , que produce una tercera función $(f * g)$ que expresa como la forma de una o bien la forma de la otra.

$$(f * g) \equiv \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau$$
$$\equiv \int_{-\infty}^{\infty} f(t - \tau) g(\tau) d\tau$$



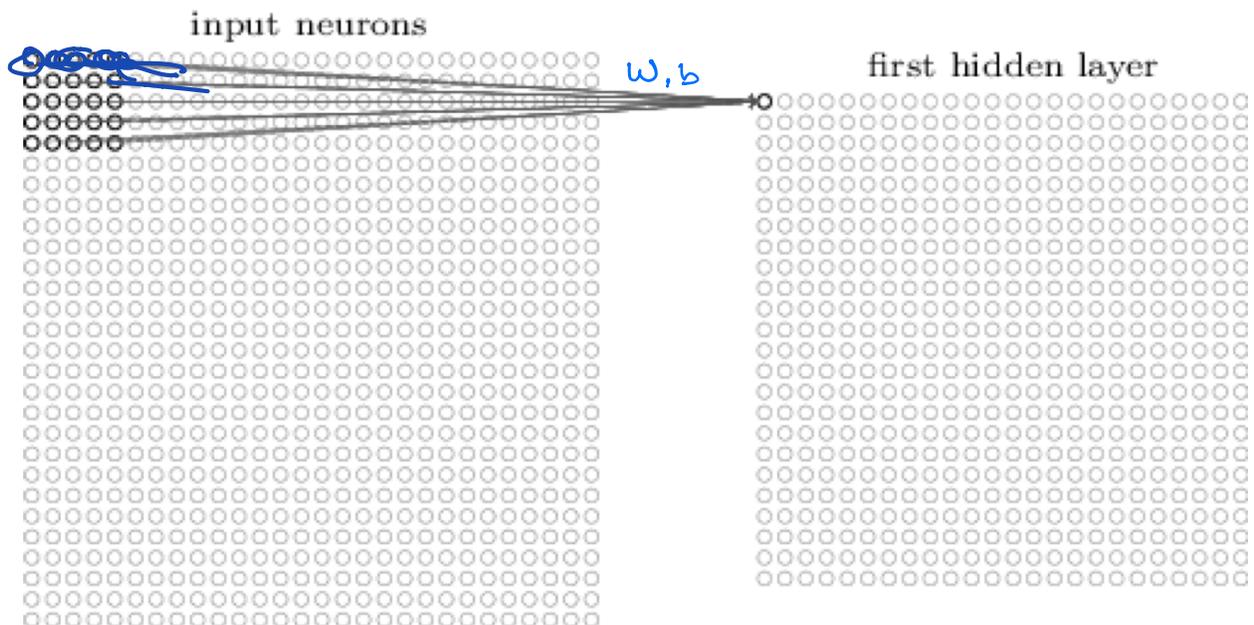
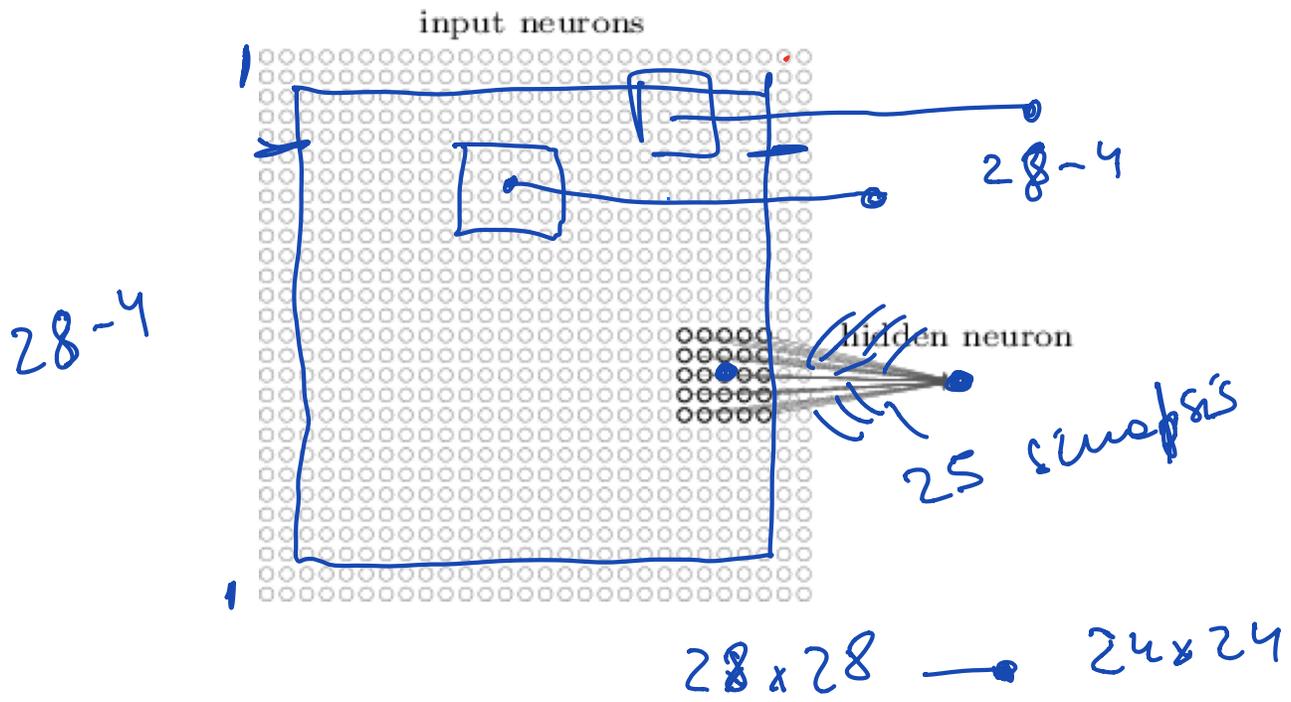
Nosotros generalizemos esta idea para definir una neurona convolucional discreta que actúe en un campo receptivo bidimensional:

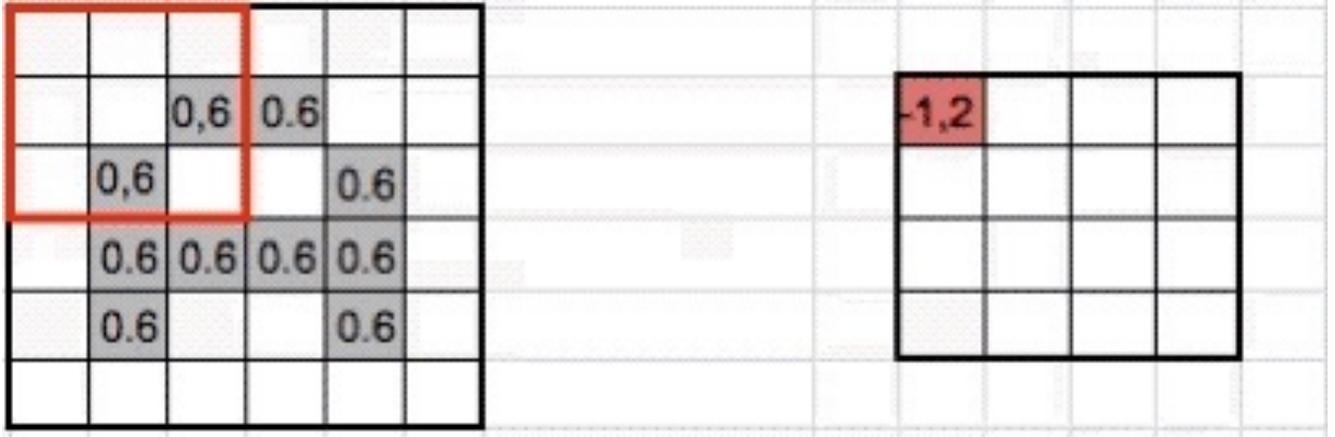
$$C[m, n] = \sum_{u=-l}^l \sum_{v=-l}^l \underbrace{W[u, v]}_{\text{kernel}} \cdot I[m-u, n-v]$$



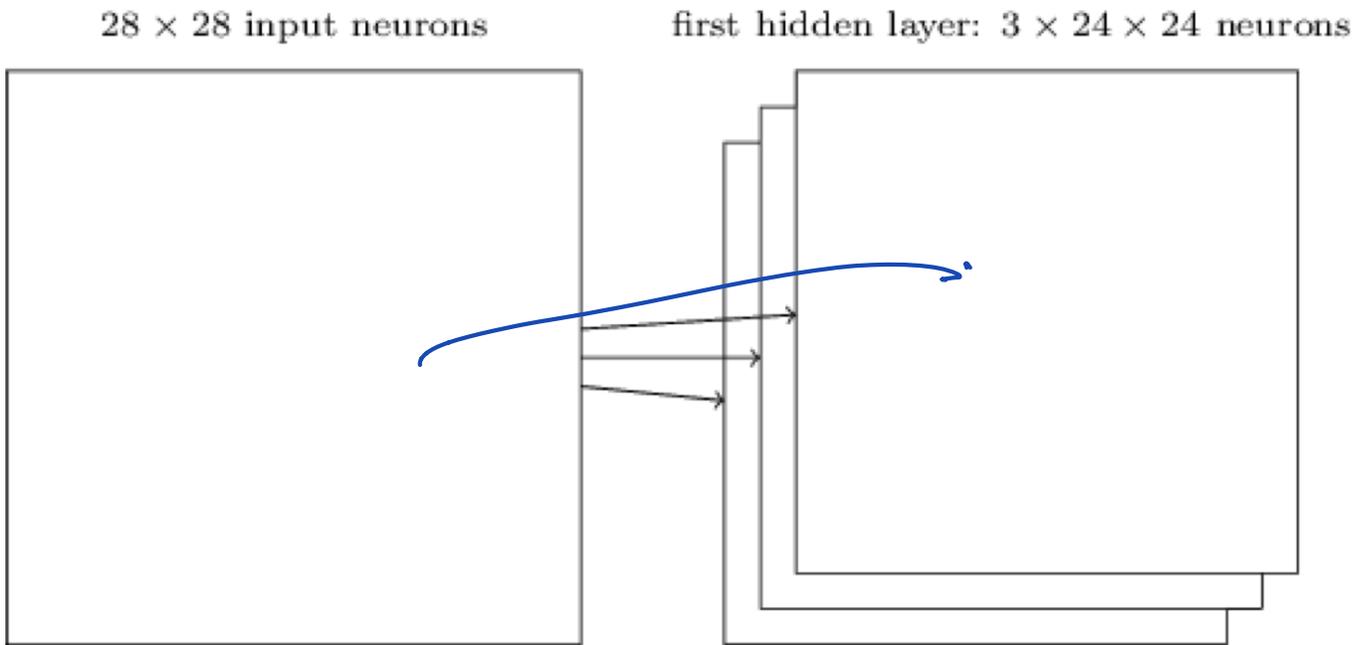
$$O[m, n] = g(C[m, n] + b)$$

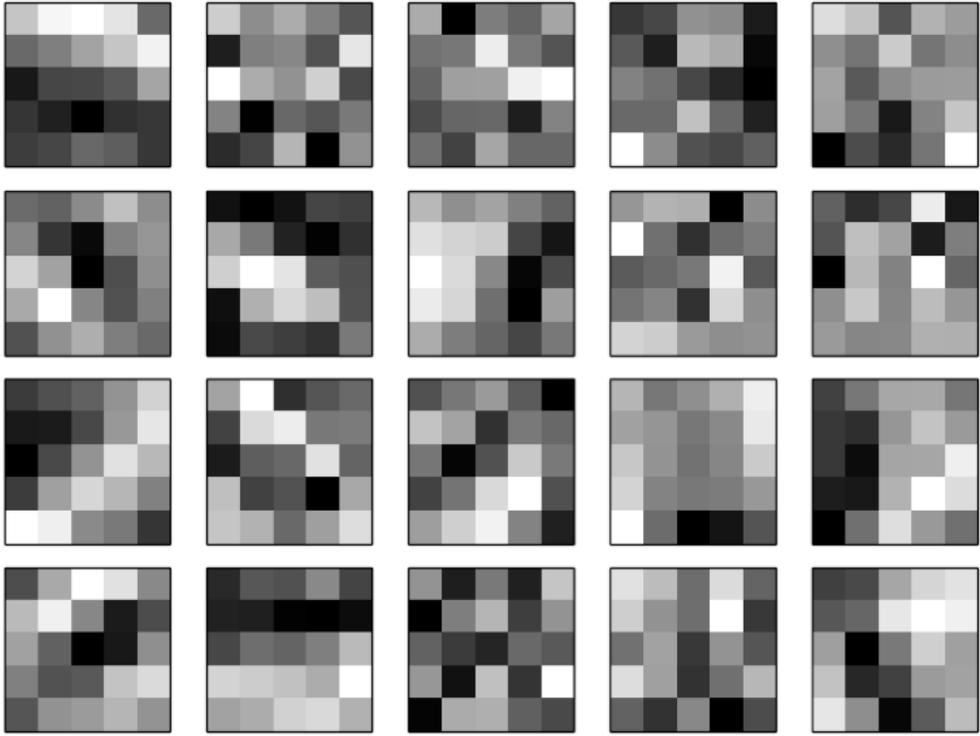
W es una matriz 2D de tamaño $(2l+1) \times (2l+1)$
 Si la entrada tiene tamaño $d \times d$, la salida tiene
 tamaño $(d-2l) \times (d-2l)$.





La red debe aprender W [m, n] y b , que es el mismo para cada filtro.

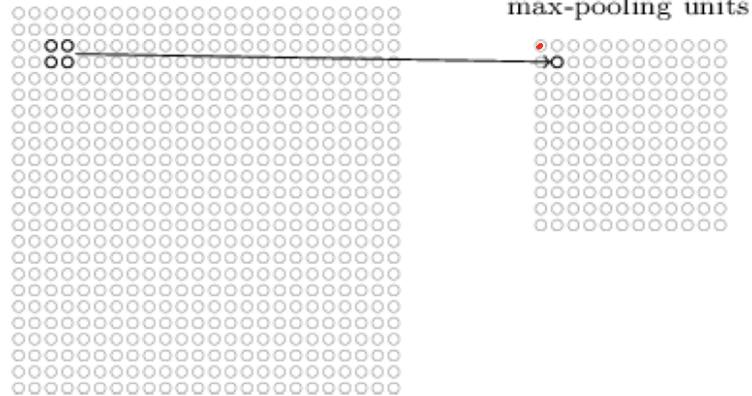




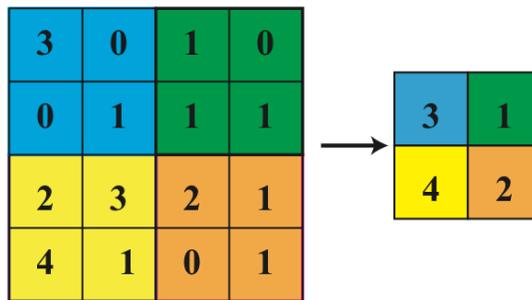
CAPA POOLING

Este caso toma el resultado de la convolución y lo reduce

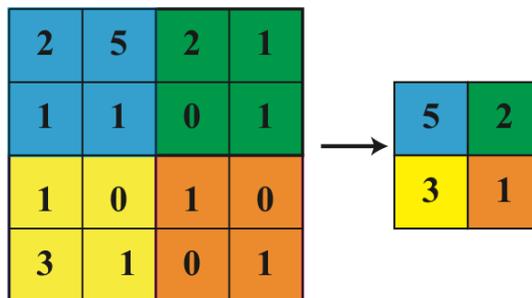
hidden neurons (output from feature map)

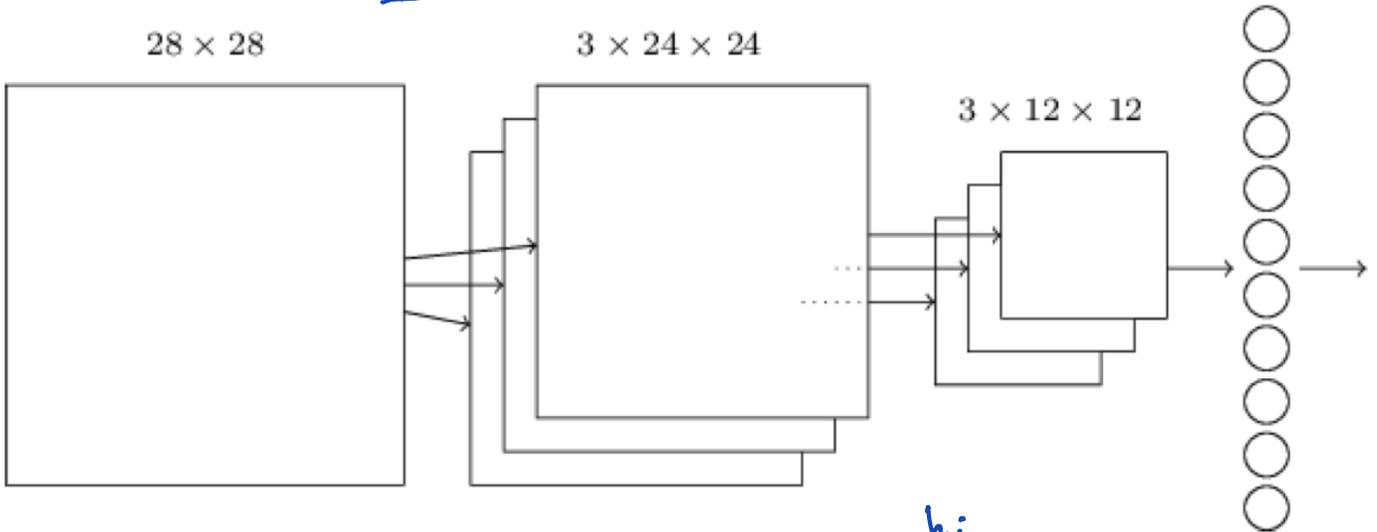
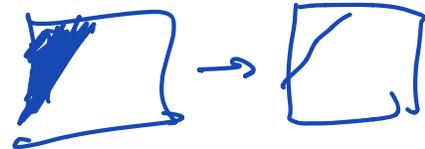
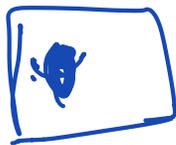
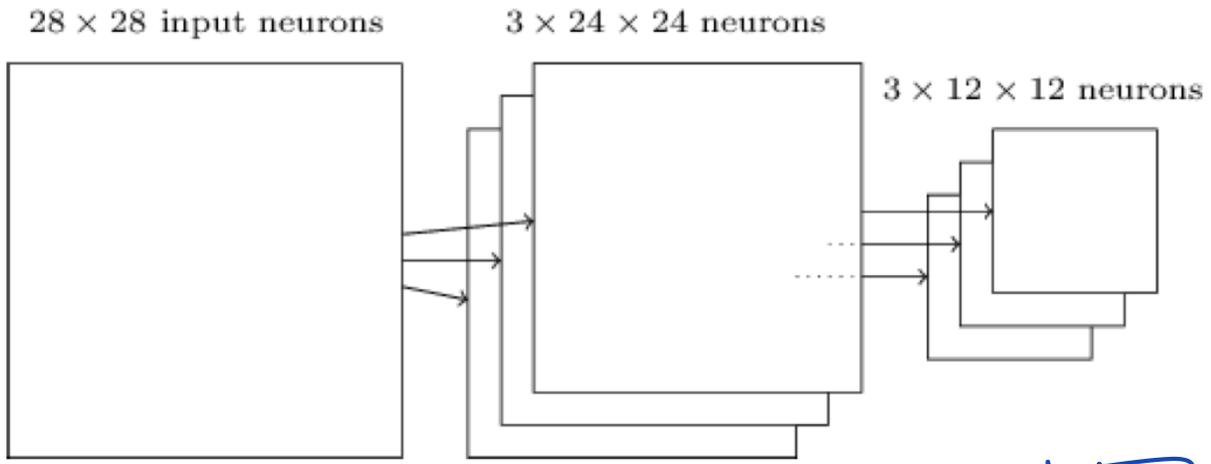


MAX POOLING



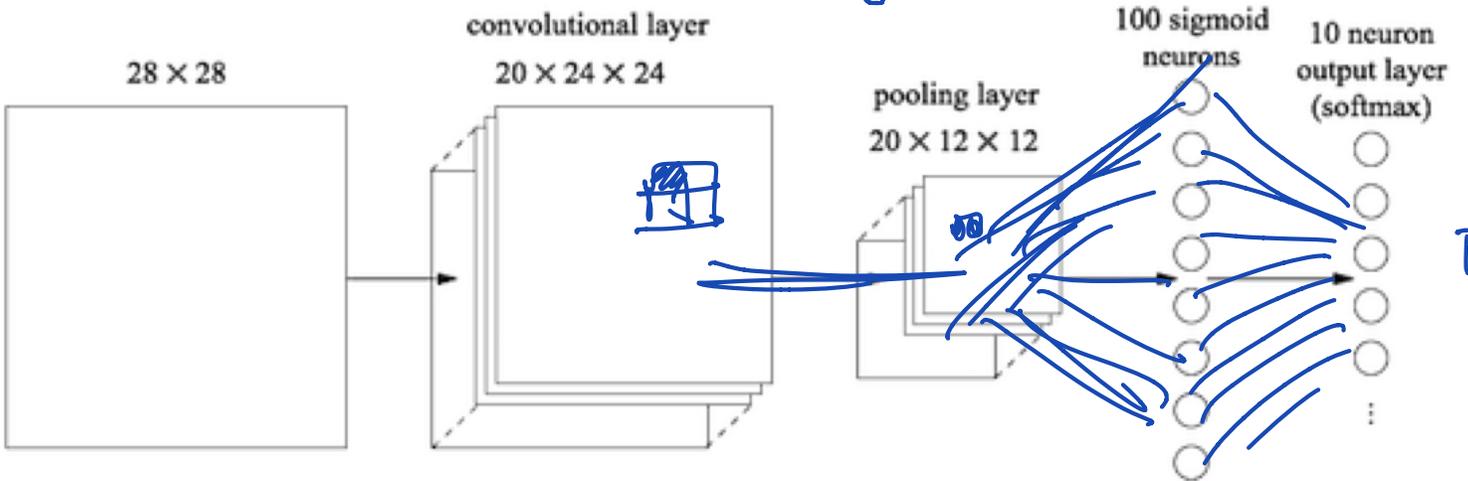
⋮ → ⋮





25 × 20

$$\text{softmax}(h_i) = \frac{e^{-h_i}}{\sum_j e^{-h_j}}$$

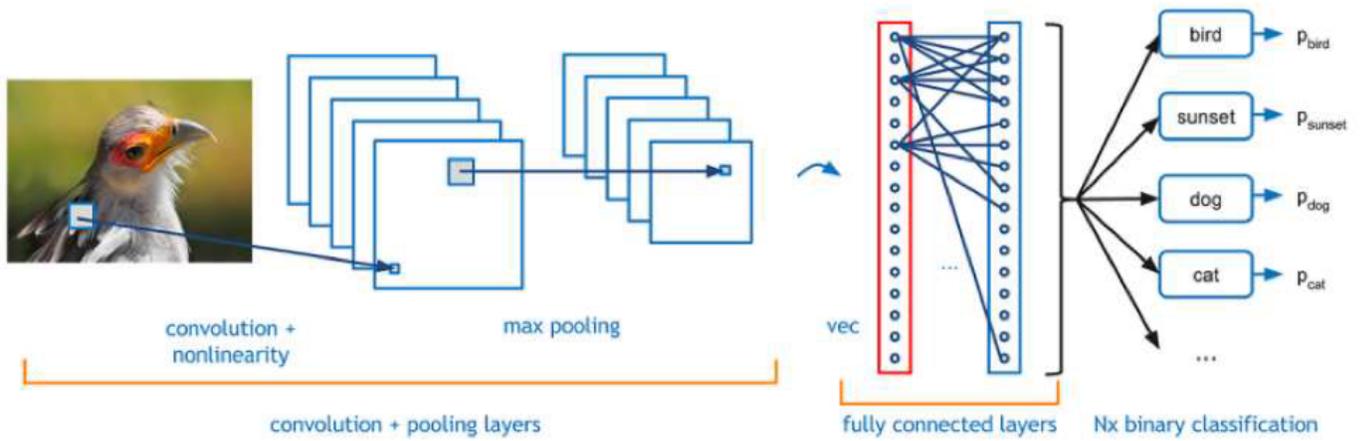
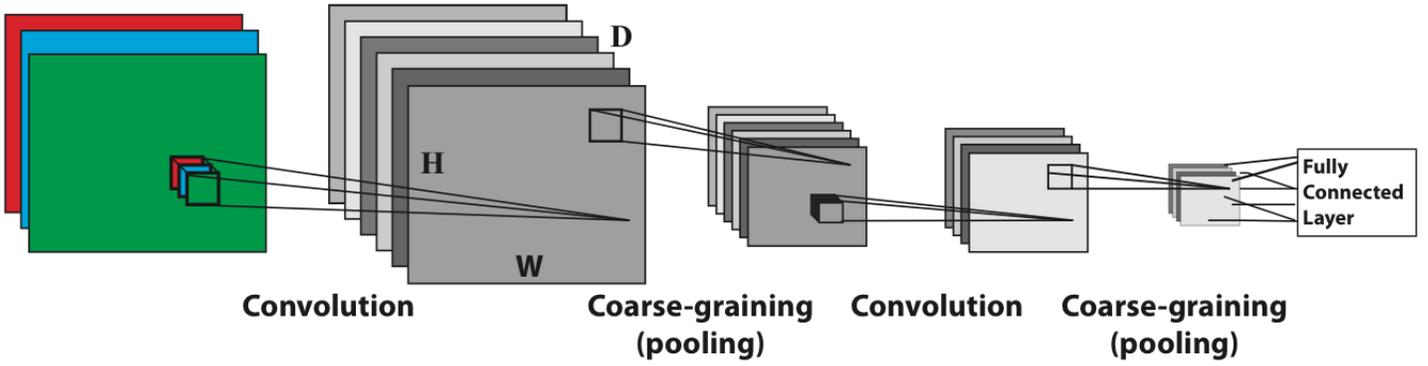


S

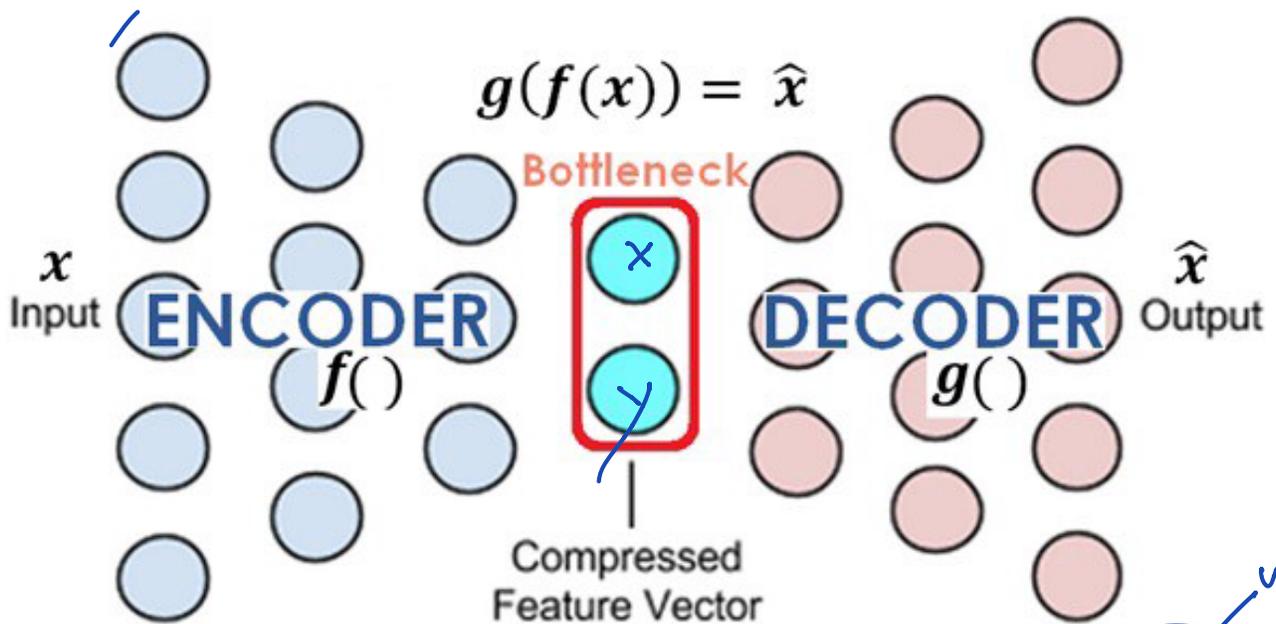
98,78%

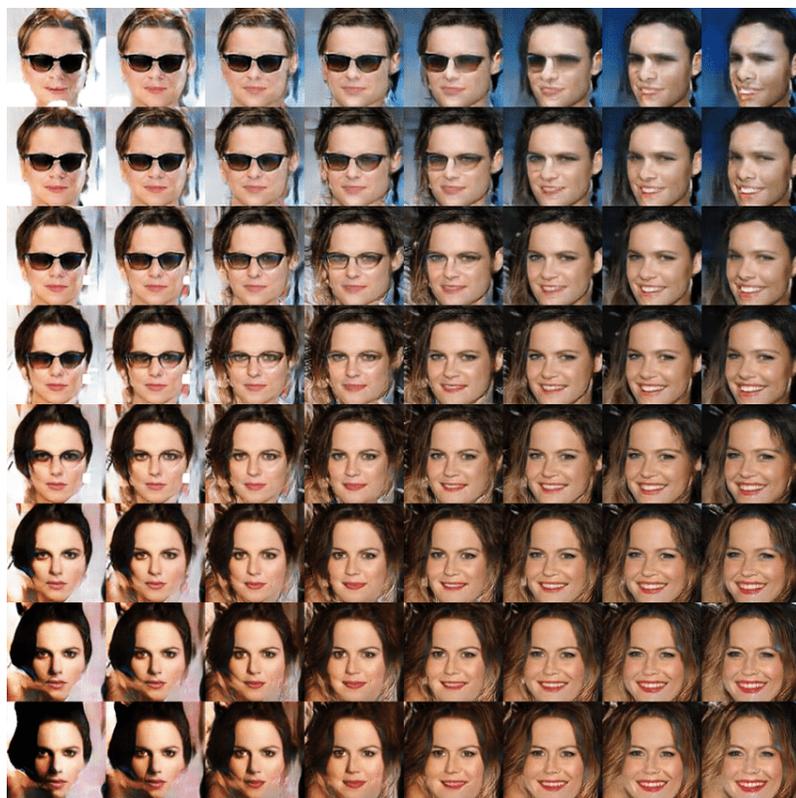
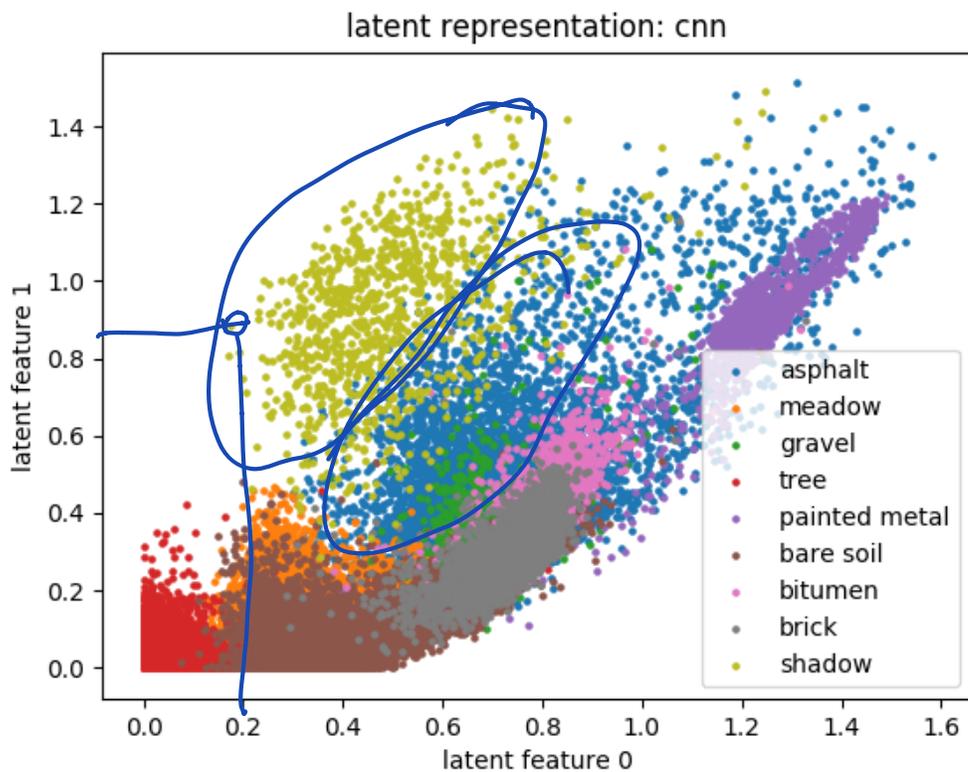
W = W

11



Autoencoder convolutional



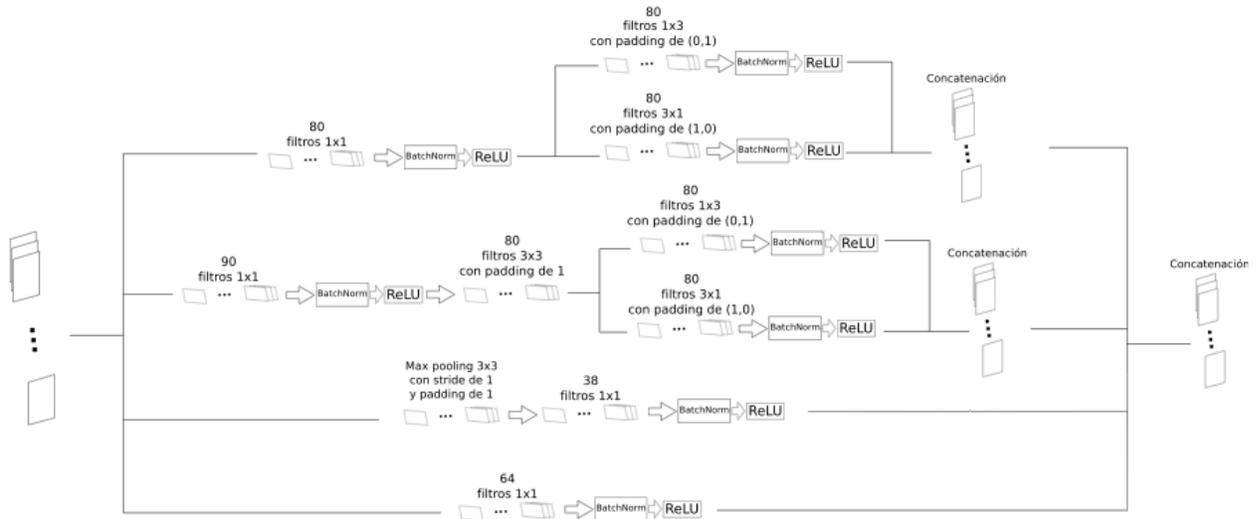




Una imagen...

| | | | | |
|-----|-----|-----|-----|--|
| | | | | |
| | 0.6 | 0.6 | | |
| 0.6 | | | 0.6 | |
| 0.6 | 0.6 | 0.6 | 0.6 | |
| 0.6 | | | 0.6 | |
| | | | | |

...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



(C) Versión modificada de la capa InceptionE implementada en PyTorch. El objetivo de esta capa es combinar información de distintos campos receptivos.

FIGURA 3.3: Capas componentes del encoder en la arquitectura Autoencoder inception.