

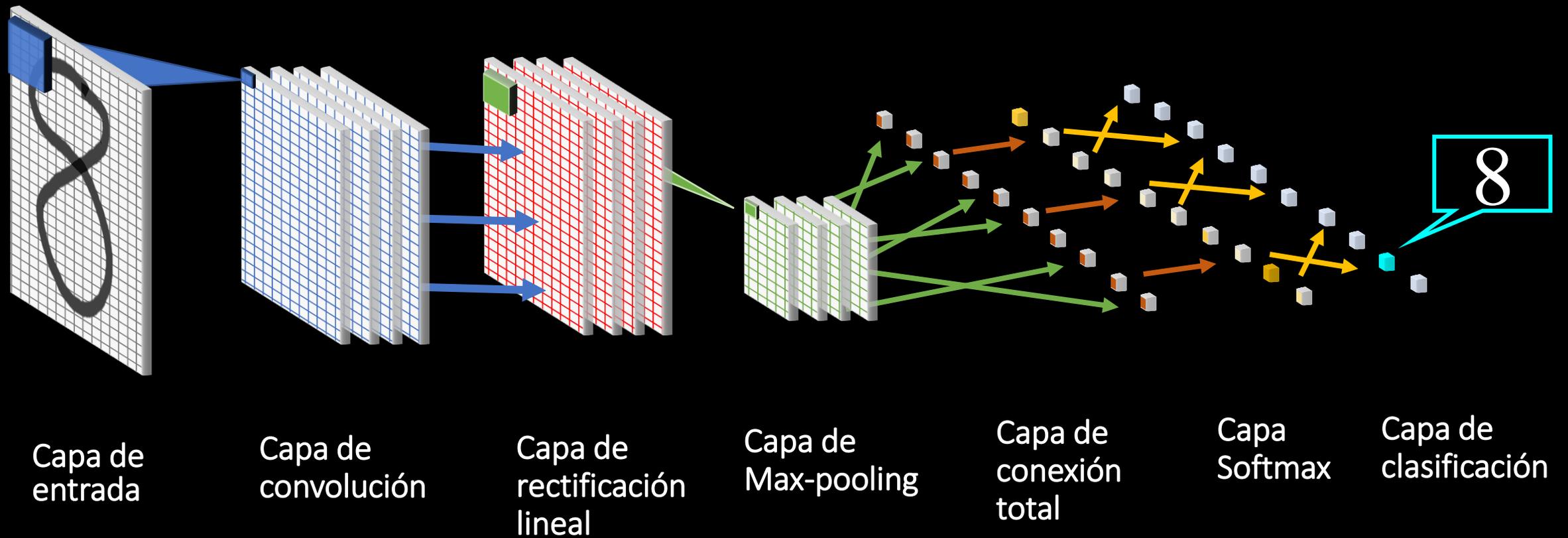
1.

Arquitectura de red convolucional Janet para el reconocimiento de dígitos manuscritos

Curso 2020-2021

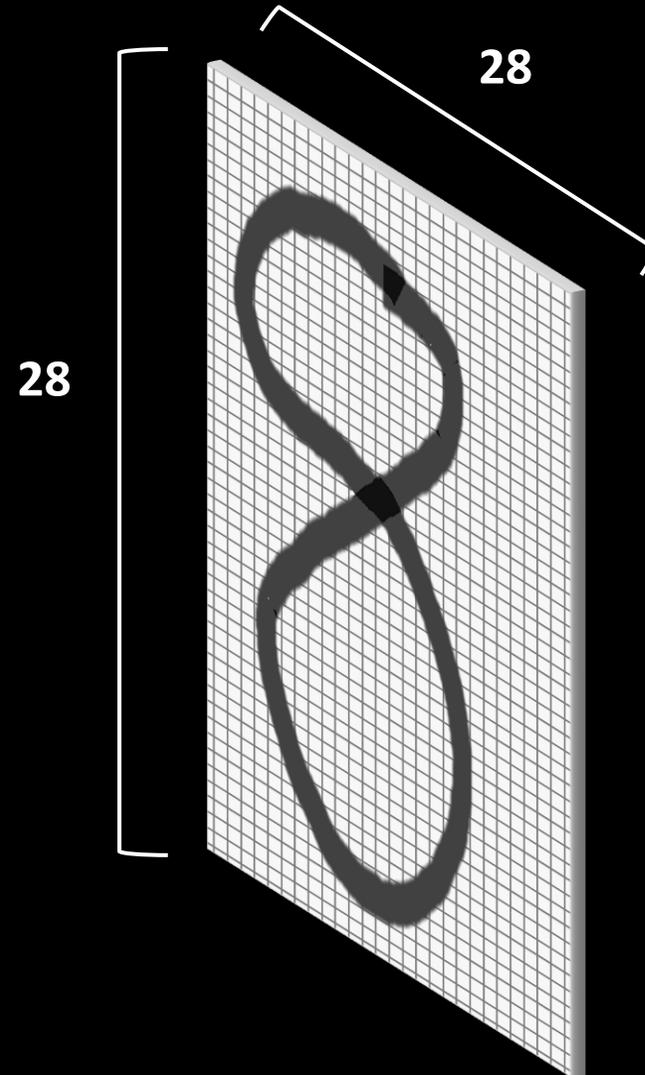
Diego Hernández Jiménez

Arquitectura completa



Capa de entrada (*input layer*)

- No se aplica ningún tipo de normalización
- No se aplica ninguna técnica de aumento de datos (*data augmentation*)
- Dimensión del output: 28 x 28 x 1

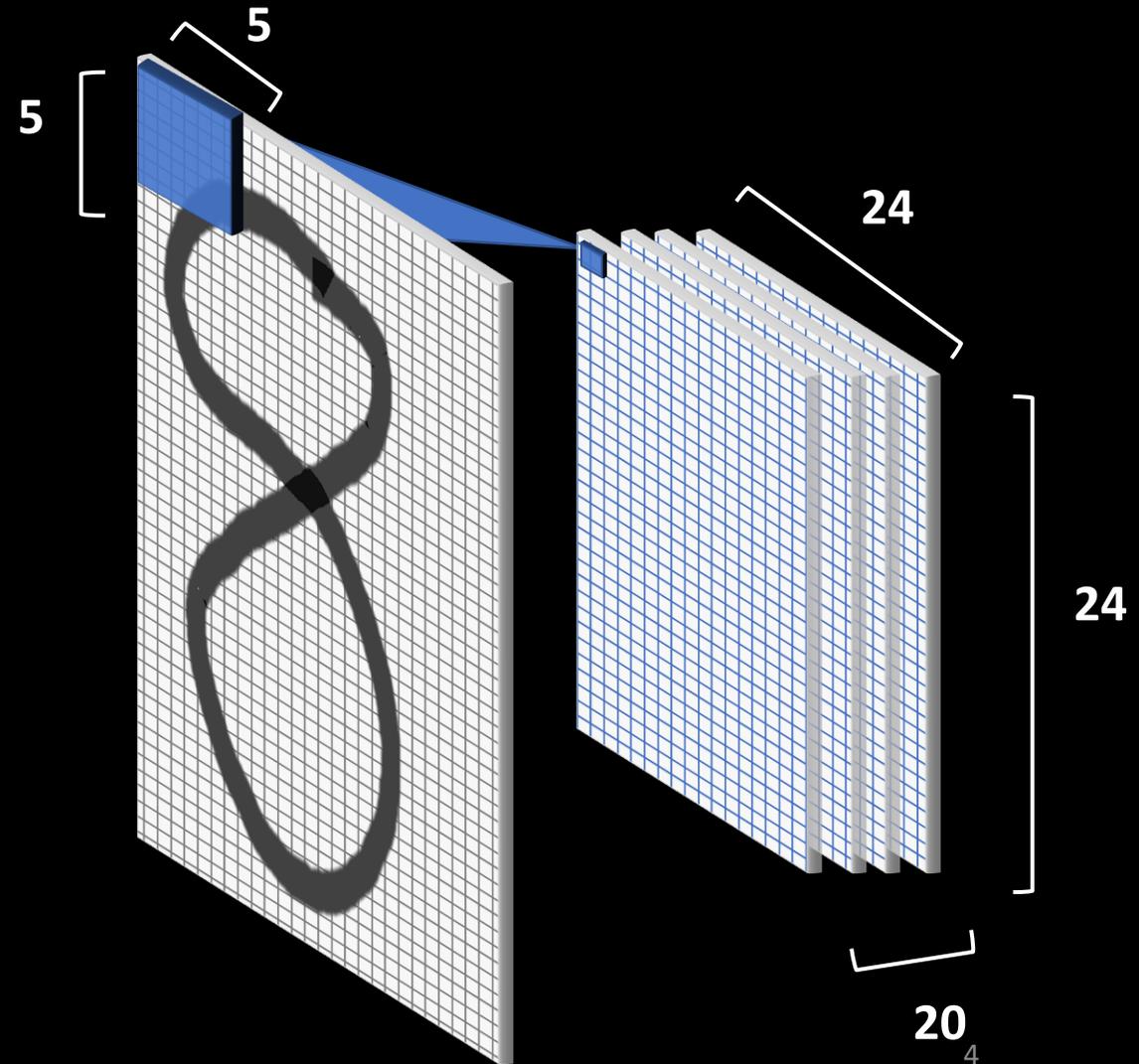


Capa convolucional (*convolution layer*)

- Dimensión del input: 28 x 28 x 1
- Nº de filtros/kernels: 20
- Nº de canales: 1
- Dimensiones del **kernel**: 5 x 5
- Paso (*stride*): [1 1]
- Relleno (*padding*): [0 0 0 0]
- Dimensión del output: 24 x 24 x 20

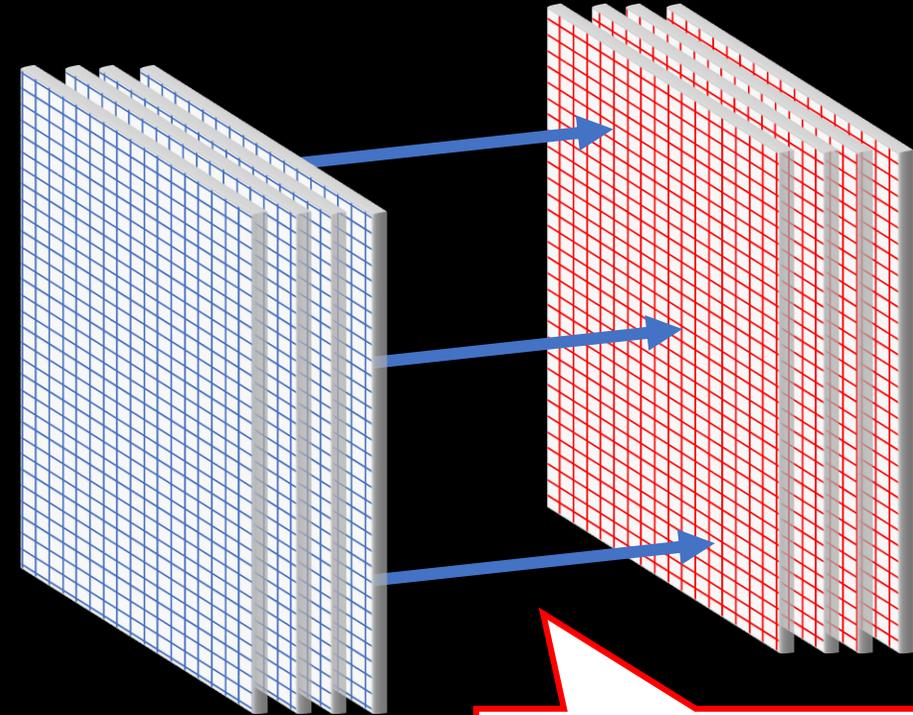
Nº parámetros:

- Pesos (w_{ij}): $20 \times 5 \times 5 = 500$
- Sesgos (b_{ij}): $20 \times 1 = 20$

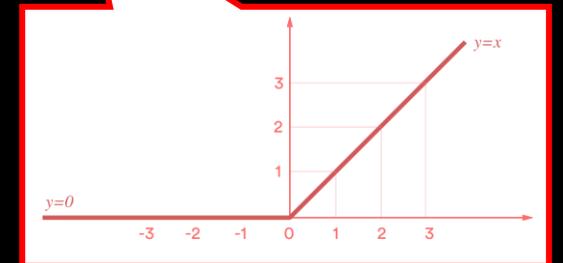


Capa de rectificación lineal (*ReLU layer*)

- Dimensión del input: 24 x 24 x 20
- Función de activación ReLU:
$$\sigma(\mathbf{Z})_{ij} = \max(0, Z_{ij})$$
- Dimensión del output: 24 x 24 x 20

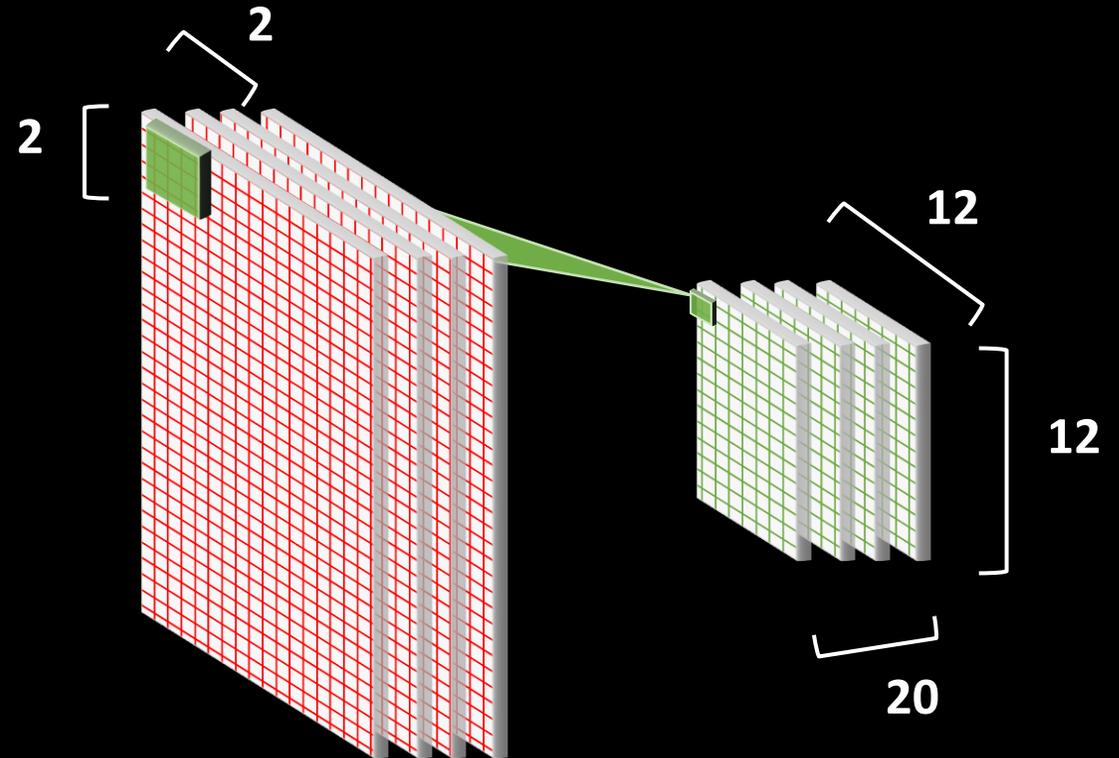


Nº parámetros: Permanece igual



Capa de Max-pooling (*Max-pooling layer*)

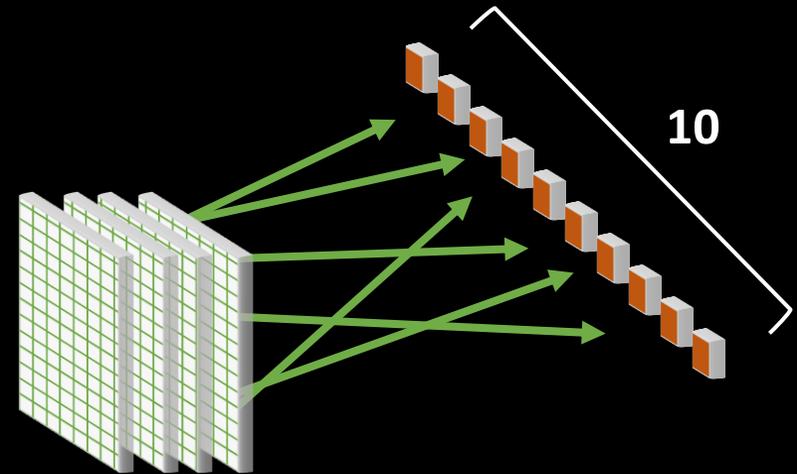
- Dimensión del input: 24 x 24 x 20
- Nº de filtros/kernels: 20
- Dimensiones del **kernel**: 2 x 2
- Paso (*stride*): [2 2]
- Relleno (*padding*): [0 0 0 0]
- Dimensión del output: 12 x 12 x 20



Nº parámetros: Permanece igual

Capa de conexión total (*Fully connected layer*)

- Dimensión del input: 12 x 12 x 20
- Nº de neuronas: 10
- Dimensión del output: 10



Nº parámetros:

- Pesos (w_{ij}): $12 \times 12 \times 20 \times 10 = 28800$
- Sesgos (b_{ij}): 10

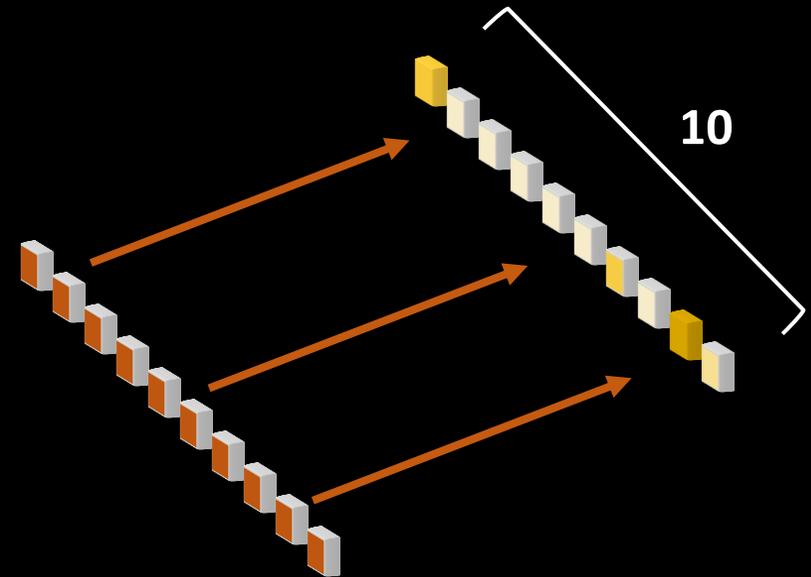
Capa Softmax (*Softmax layer*)

- Dimensión del input: 10

- Función softmax:

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}}$$

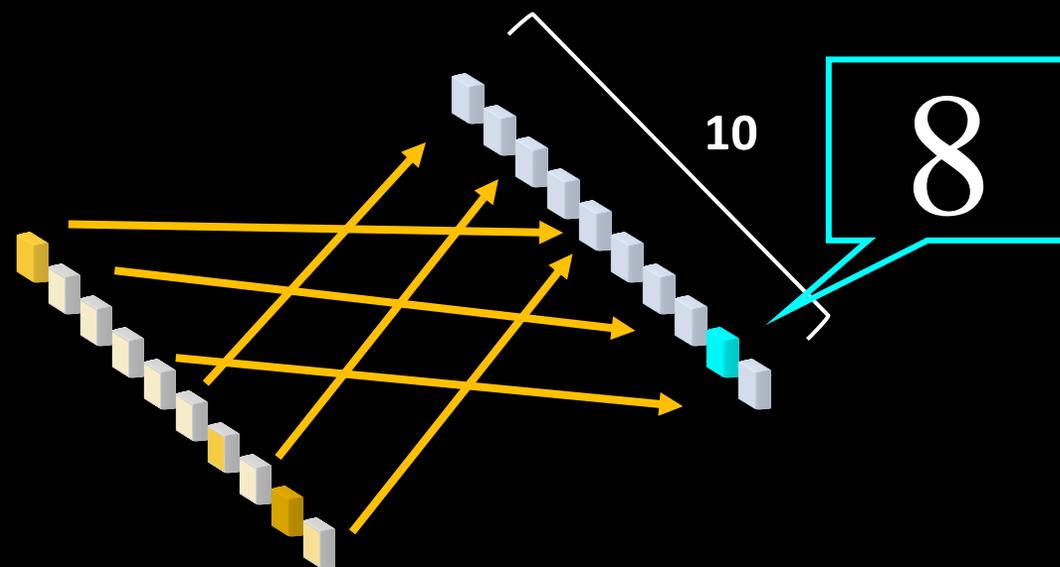
- Dimensión del output: 10



Nº parámetros: Permanece igual

Capa de clasificación (*classification layer*)

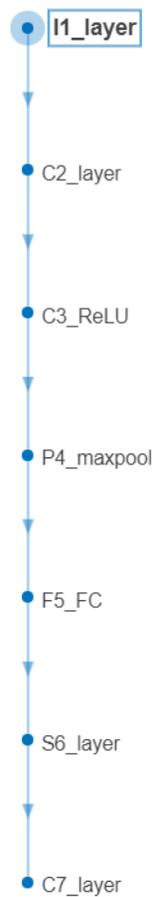
- Dimensión del input: 10
- Dimensión del output: 10



Nº parámetros: Permanece igual

Número total de parámetros involucrados

	<u>Pesos + sesgos</u>	
• Capa de entrada	0	
• Capa de convolución	500 + 20	
• Capa de rectificación lineal	0	
• Capa de Max-pooling	0	
• Capa de conexión total	28800 + 10	
• Capa Softmax	0	
• Capa de clasificación	0	
		= 29330



ANALYSIS RESULT

	Name	Type	Activations	Learnables
1	I1_layer 28×28×1 images	Image Input	28×28×1	-
2	C2_layer 20 5×5×1 convolutions with stride [1 1] and padding [0 0 0 0]	Convolution	24×24×20	Weights 5×5×1×20 Bias 1×1×20
3	C3_ReLU ReLU	ReLU	24×24×20	-
4	P4_maxpool 2×2 max pooling with stride [2 2] and padding [0 0 0 0]	Max Pooling	12×12×20	-
5	F5_FC 10 fully connected layer	Fully Connected	1×1×10	Weights 10×2880 Bias 10×1
6	S6_layer softmax	Softmax	1×1×10	-
7	C7_layer crossentropyex	Classification Output	-	-