

Aprendizaje Automático

Clase 22: Redes Neuronales Convolucionales

Historia de las CNN (Convolutional Neural Nets)

- * En 1995, Yann LeCun y Yoshua Bengio introdujeron el concepto de redes neuronales convolucionales o CNN del inglés “Convolutional Neural Nets”.
- * Son redes que se usan para procesar imágenes
- * Pueden aprender relaciones entrada-salida, donde la entrada es una imagen
- * Están basadas en operaciones de convolución
- * Tareas comunes:
 - * Detección/categorización de objetos
 - * Clasificación de escenas
 - * Clasificación de imágenes en general

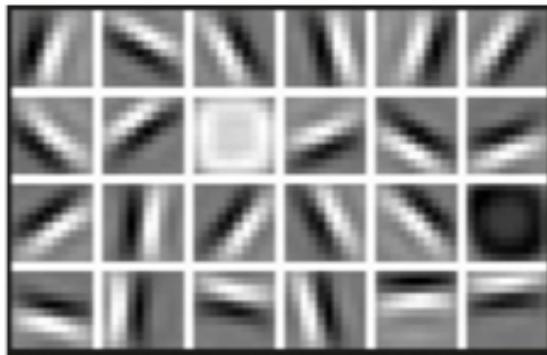
Acerca de las Redes Neuronales Convolucionales

- * CNN's fueron neurobiológicamente motivadas por los hallazgos de células nerviosas localmente sensibles y selectivas de la orientación en el visual corteza.
- * Diseñaron una estructura de red que implícitamente extrae características relevantes.
- * Las redes neuronales convolucionales son un tipo especial de Redes neuronales multicapa.

Acerca de las Redes Neuronales Convolucionales

- * CNN es una red de propagación hacia adelante que puede extraer propiedades topológicas de una imagen.
- * Como casi cualquier otra red neuronal, se entrena con una versión del algoritmo de propagación hacia atrás.
- * Las redes neuronales convolucionales están diseñadas para reconocer patrones directamente desde imágenes de píxeles con un preprocessamiento mínimo.
- * Pueden reconocer patrones con variabilidad extrema (como caracteres escritos a mano).

Redes Neuronales Convolucionales



First Layer Representation



Second Layer Representation

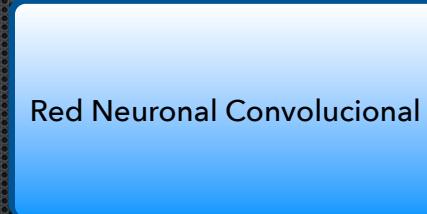
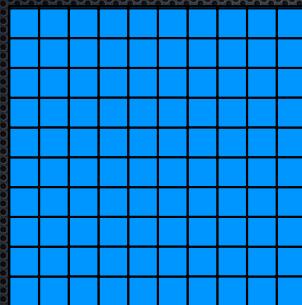


Third Layer Representation

Un Ejemplo Sencillo de CNN

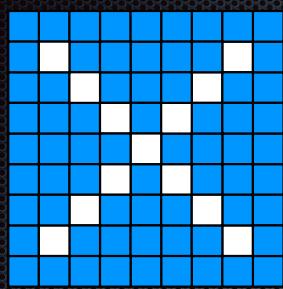
- * Ejemplo de una CNN para reconocer si una imagen es una X o una O

Un arreglo de pixeles
de dos dimensiones



X
O
0

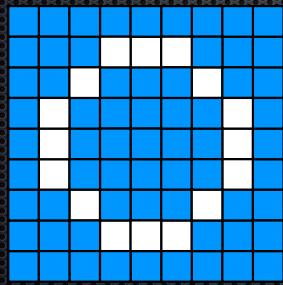
Por ejemplo:



Red Neuronal Convolucional



X

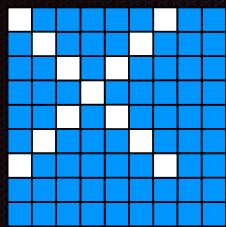


Red Neuronal Convolucional

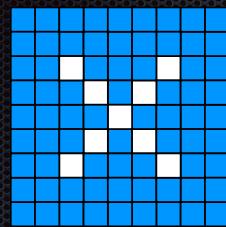


0

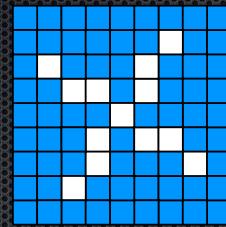
Casos difíciles



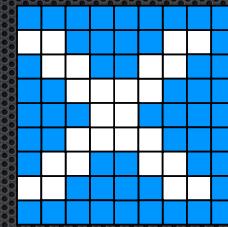
Traslación



Escalamiento



Rotación



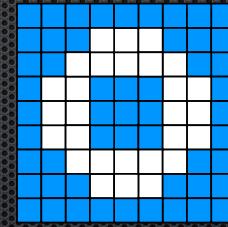
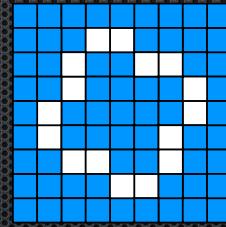
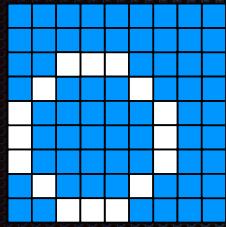
Grosor



Red Neuronal Convolucional



X

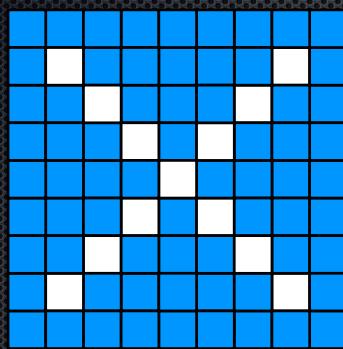


Red Neuronal Convolucional

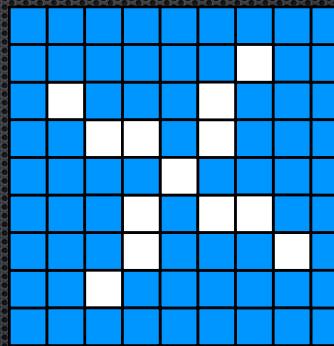


0

Decidir es difícil



?
=



¿Qué es lo que ve el computador?

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

?

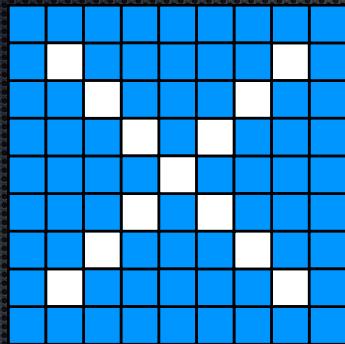
=

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	1	-1
-1	1	-1	-1	-1	-1	1	-1	-1
-1	-1	1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	1	-1	-1	-1
-1	-1	-1	-1	-1	1	-1	-1	-1
-1	-1	-1	1	-1	1	1	-1	-1
-1	-1	-1	1	-1	-1	-1	-1	1
-1	-1	-1	-1	-1	-1	-1	-1	-1

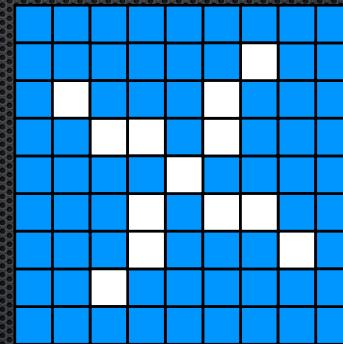
¿Qué es lo que ve el computador?

-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	X	-1	-1	-1	-1	X	X	-1	
-1	X	X	-1	-1	X	X	-1	-1	
-1	-1	X	1	-1	1	-1	-1	-1	
-1	-1	-1	-1	1	-1	-1	-1	-1	
-1	-1	-1	1	-1	1	X	-1	-1	
-1	-1	X	X	-1	-1	X	X	-1	
-1	X	X	-1	-1	-1	-1	X	-1	
-1	-1	-1	-1	-1	-1	-1	-1	-1	

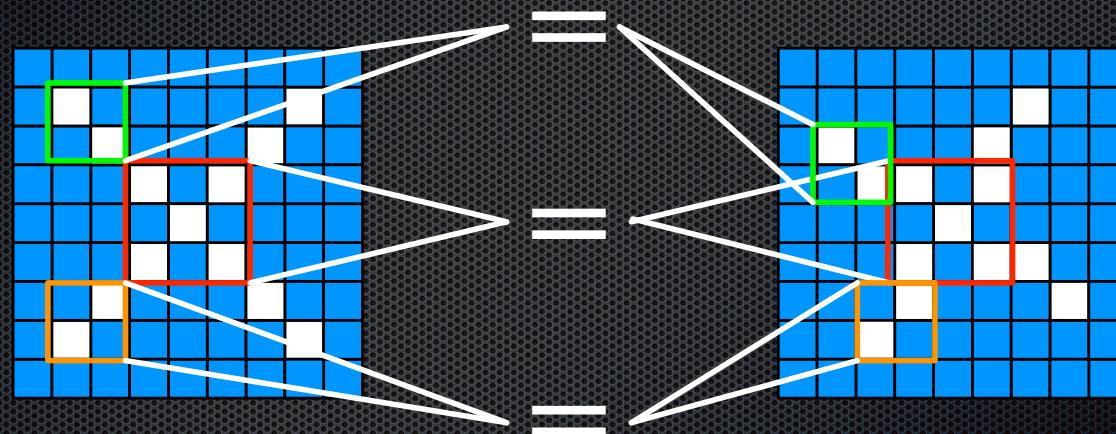
Computador decide que son diferentes



✗



Redes Neuronales Convolucionales emparejan partes de la imagen



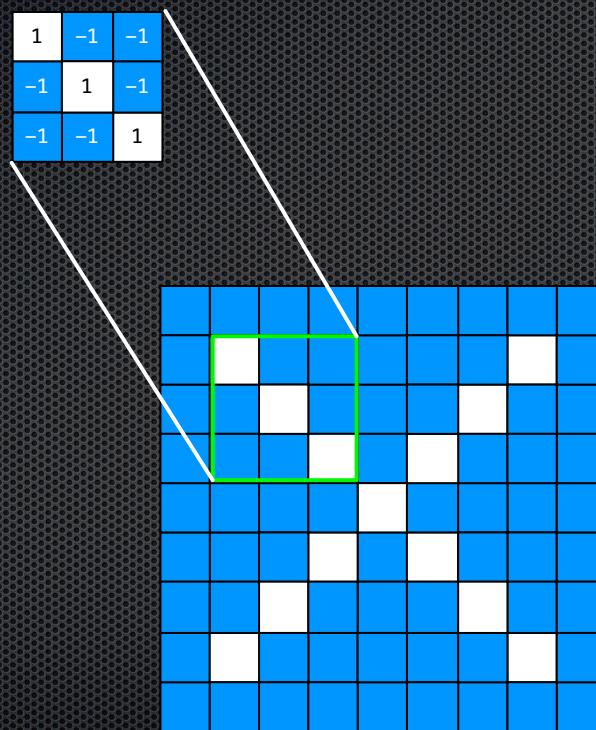
Filtros emparejan pedazos de la imagen

1	-1	-1
-1	1	-1
-1	-1	1

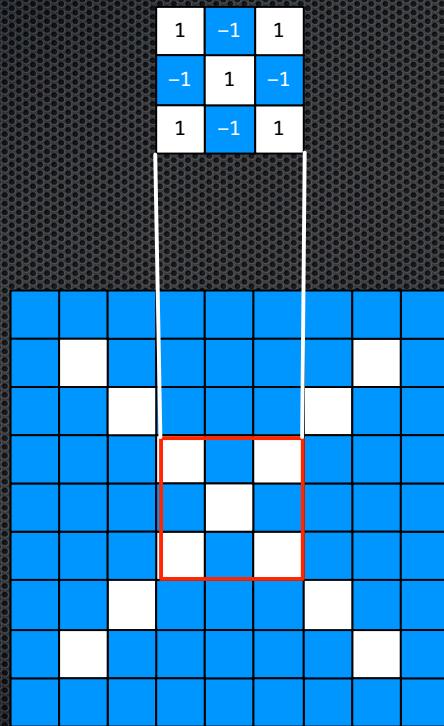
1	-1	1
-1	1	-1
1	-1	1

-1	-1	1
-1	1	-1
1	-1	-1

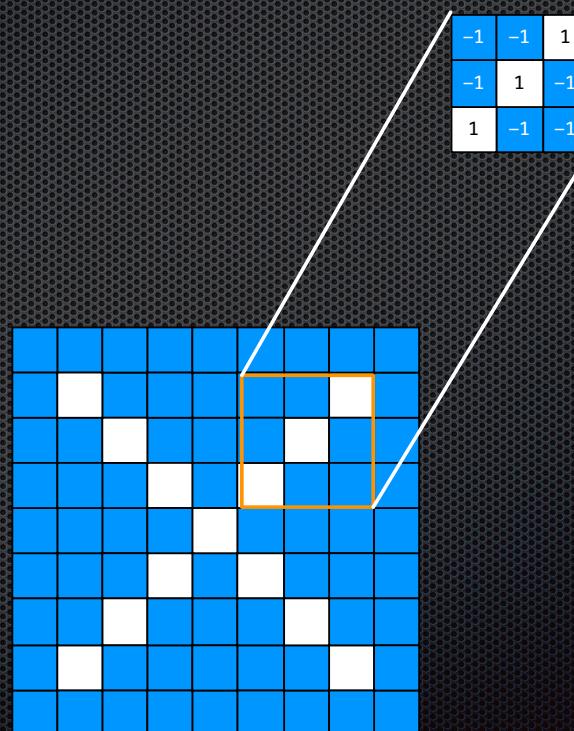
Filtro para línea diagonal 135°



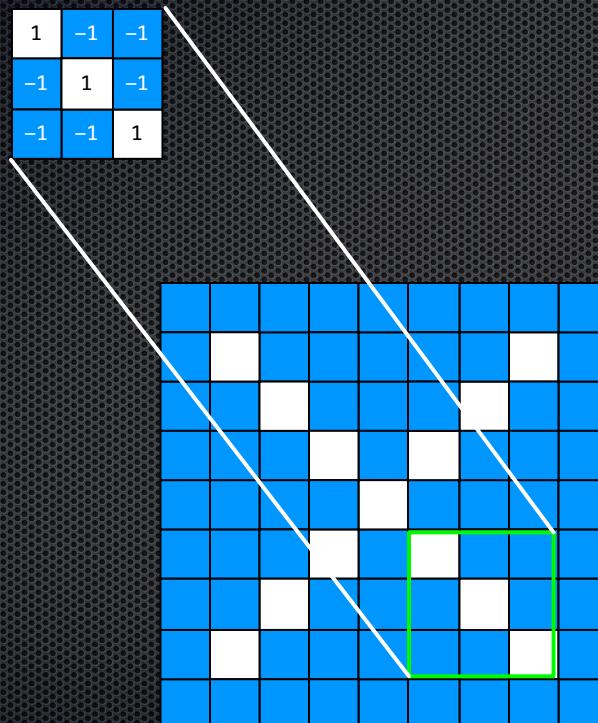
Filtro para la X



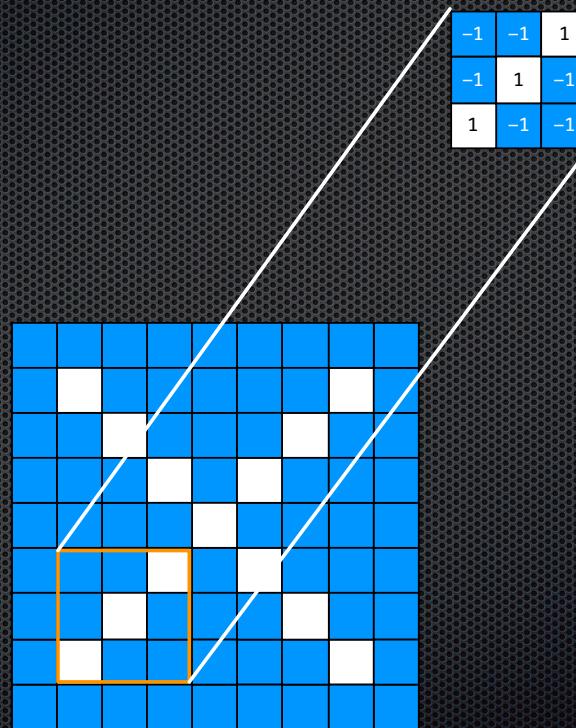
Filtro para línea diagonal 45°



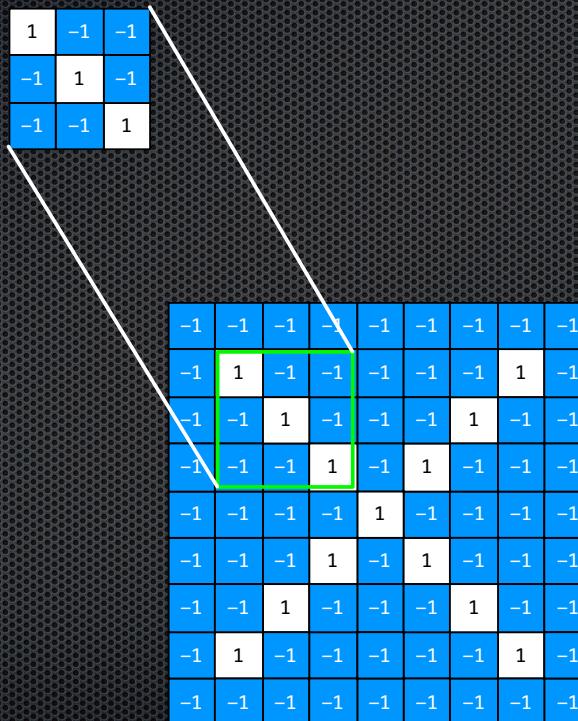
Filtro para línea diagonal 135°



Filtro para línea diagonal 45°



Matemáticas del Filtrado



Matemáticas del Filtrado

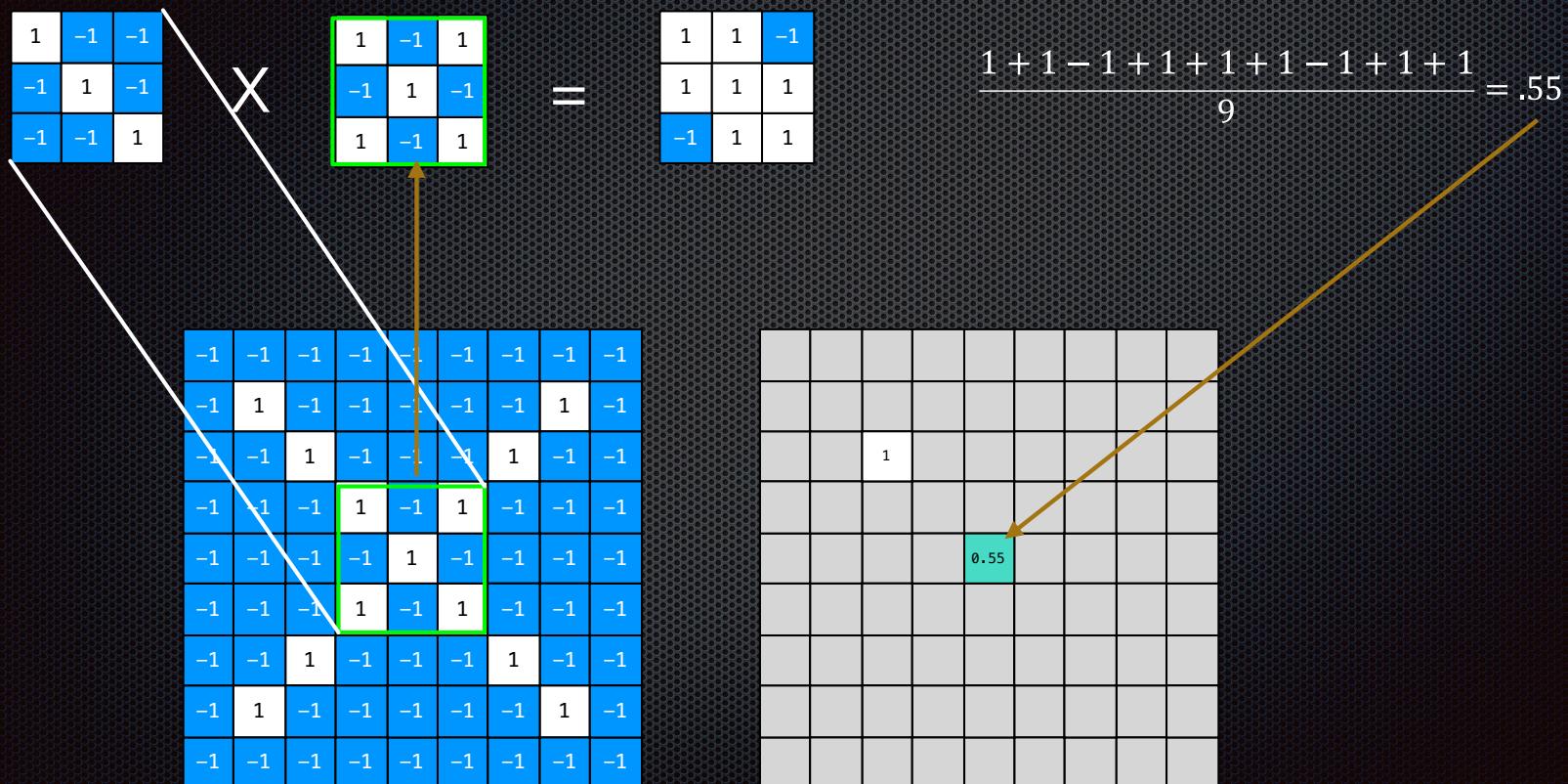
- * Alinee el filtro y el parche de la imagen.
- * Multiplique cada píxel de la imagen con el píxel correspondiente del filtro.
- * Sumelos.
- * Divida por el número total de píxeles en el filtro

Matemáticas del Filtrado

$$\begin{array}{c} \begin{matrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{matrix} \times \begin{matrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{matrix} = \begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix} \\ \frac{1 + 1 + 1 + 1 + 1 + 1 + 1 + 1}{9} = 1 \end{array}$$

The diagram illustrates a convolution operation. On the left, two 3x3 matrices are shown: a blue input matrix and a green kernel matrix. The multiplication results in a 3x3 output matrix where every element is 1. To the right of the multiplication, a mathematical expression shows the sum of all elements in the output matrix divided by 9, resulting in 1. Below the input matrix, a 9x9 grid represents the receptive field of the central output unit. The central cell of this grid is highlighted in green, and an arrow points from the central output unit to this green cell.

Matemáticas del Filtrado



Resultado Final de la Convolución con Filtro 135°

$$\begin{array}{|c|c|c|c|c|c|c|c|c|c|} \hline -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ \hline -1 & 1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 \\ \hline -1 & -1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 \\ \hline -1 & -1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 \\ \hline -1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 & -1 \\ \hline -1 & -1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 \\ \hline -1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 & -1 \\ \hline -1 & -1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 \\ \hline -1 & 1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 \\ \hline -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ \hline \end{array} \otimes \begin{array}{|c|c|c|} \hline 1 & -1 & -1 \\ \hline -1 & 1 & -1 \\ \hline -1 & -1 & 1 \\ \hline \end{array} = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.77 & -0.11 & 0.11 & 0.33 & 0.55 & -0.11 & 0.33 \\ \hline -0.11 & 1.00 & -0.11 & 0.33 & -0.11 & 0.11 & -0.11 \\ \hline 0.11 & -0.11 & 1.00 & -0.33 & 0.11 & -0.11 & 0.55 \\ \hline 0.33 & 0.33 & -0.33 & 0.55 & -0.33 & 0.33 & 0.33 \\ \hline 0.55 & -0.11 & 0.11 & -0.33 & 1.00 & -0.11 & 0.11 \\ \hline -0.11 & 0.11 & -0.11 & 0.33 & -0.11 & 1.00 & -0.11 \\ \hline 0.33 & -0.11 & 0.55 & 0.33 & 0.11 & -0.11 & 0.77 \\ \hline \end{array}$$

Resultado Final de la Convolución con Filtro para X

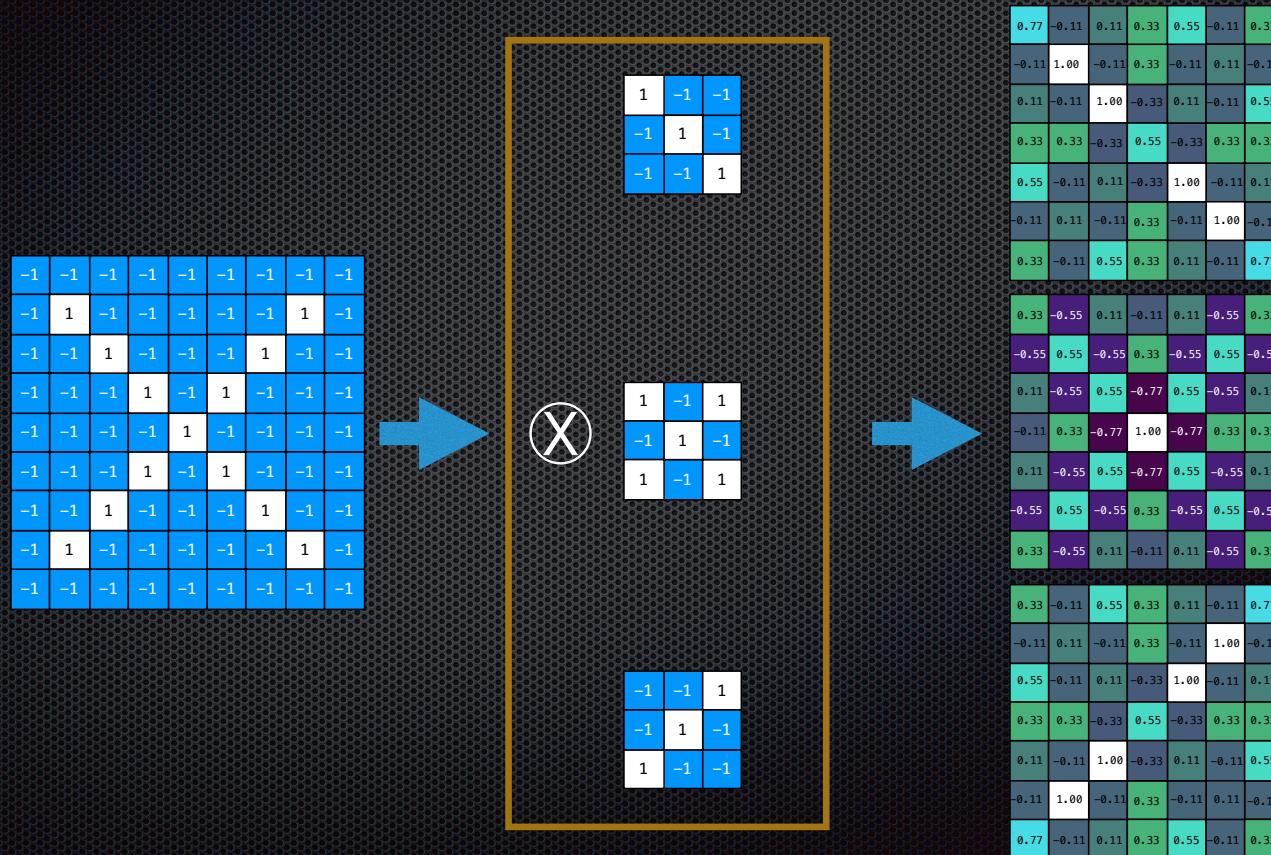
$$\begin{array}{|c|c|c|c|c|c|c|c|c|c|} \hline -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ \hline -1 & 1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 \\ \hline -1 & -1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 \\ \hline -1 & -1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 \\ \hline -1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 & -1 \\ \hline -1 & -1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 \\ \hline -1 & -1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 \\ \hline -1 & -1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 \\ \hline -1 & 1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 \\ \hline -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ \hline \end{array} \otimes \begin{array}{|c|c|c|} \hline 1 & -1 & 1 \\ \hline -1 & 1 & -1 \\ \hline 1 & -1 & 1 \\ \hline \end{array} = \begin{array}{|c|c|c|c|c|c|c|c|c|c|} \hline 0.33 & -0.55 & 0.11 & -0.11 & 0.11 & -0.55 & 0.33 \\ \hline -0.55 & 0.55 & -0.55 & 0.33 & -0.55 & 0.55 & -0.55 \\ \hline 0.11 & -0.55 & 0.55 & -0.77 & 0.55 & -0.55 & 0.11 \\ \hline -0.11 & 0.33 & -0.77 & 1.00 & -0.77 & 0.33 & 0.33 \\ \hline 0.11 & -0.55 & 0.55 & -0.77 & 0.55 & -0.55 & 0.11 \\ \hline -0.55 & 0.55 & -0.55 & 0.33 & -0.55 & 0.55 & -0.55 \\ \hline 0.33 & -0.55 & 0.11 & -0.11 & 0.11 & -0.55 & 0.33 \\ \hline \end{array}$$

Resultado Final de la Convolución con Filtro 45°

$$\begin{array}{|c|c|c|c|c|c|c|c|c|c|} \hline -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ \hline -1 & 1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 \\ \hline -1 & -1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 \\ \hline -1 & -1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 \\ \hline -1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 & -1 \\ \hline -1 & -1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 \\ \hline -1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 & -1 \\ \hline -1 & 1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 \\ \hline -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ \hline \end{array} \otimes \begin{array}{|c|c|c|} \hline -1 & -1 & 1 \\ \hline -1 & 1 & -1 \\ \hline 1 & -1 & -1 \\ \hline \end{array} = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.33 & -0.11 & 0.55 & 0.33 & 0.11 & -0.11 & 0.77 \\ \hline -0.11 & 0.11 & -0.11 & 0.33 & -0.11 & 1.00 & -0.11 \\ \hline 0.55 & -0.11 & 0.11 & -0.33 & 1.00 & -0.11 & 0.11 \\ \hline 0.33 & 0.33 & -0.33 & 0.55 & -0.33 & 0.33 & 0.33 \\ \hline 0.11 & -0.11 & 1.00 & -0.33 & 0.11 & -0.11 & 0.55 \\ \hline -0.11 & 1.00 & -0.11 & 0.33 & -0.11 & 0.11 & -0.11 \\ \hline 0.77 & -0.11 & 0.11 & 0.33 & 0.55 & -0.11 & 0.33 \\ \hline \end{array}$$

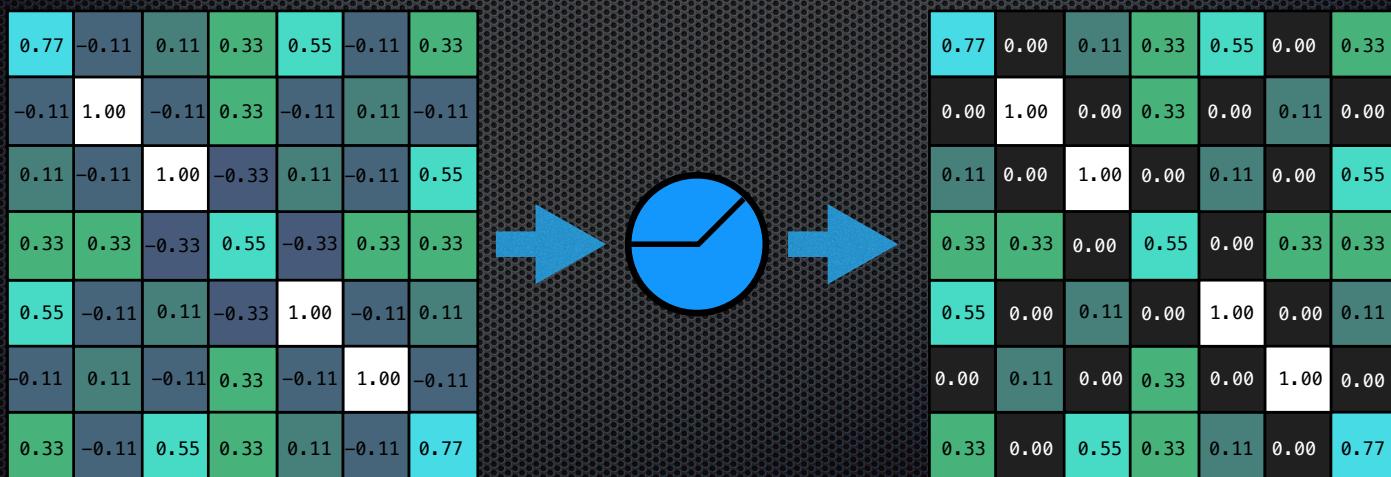
Capa de Convolución

- * Una imagen se convierte en una pila de imágenes filtradas



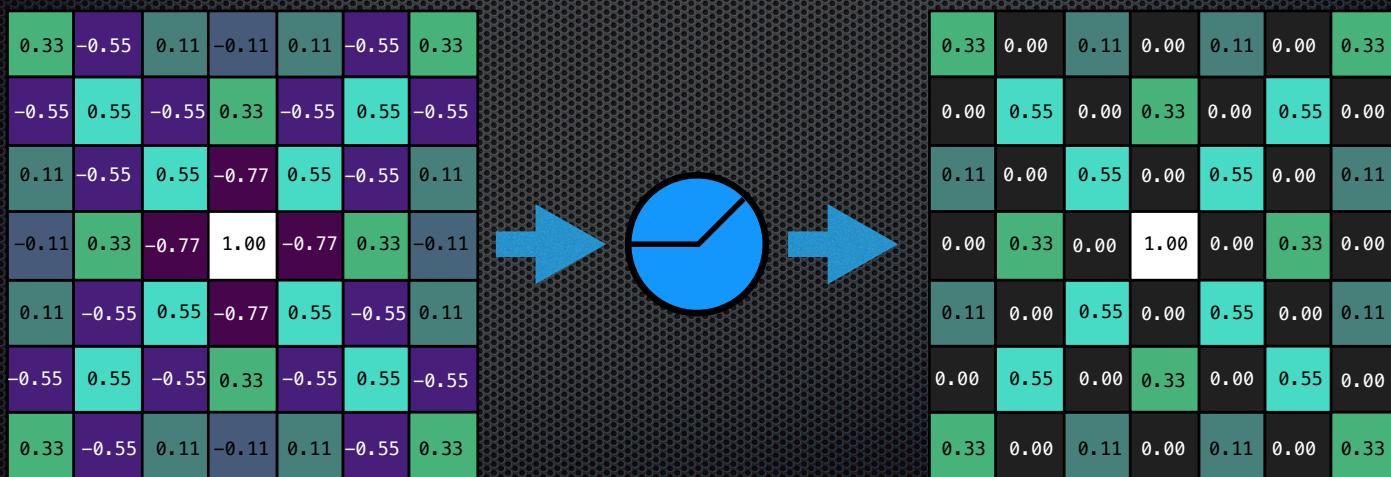
Capa ReLU:

- * Esta capa aplica la función de activación ReLU, la cual convierte todos los valores negativos a cero



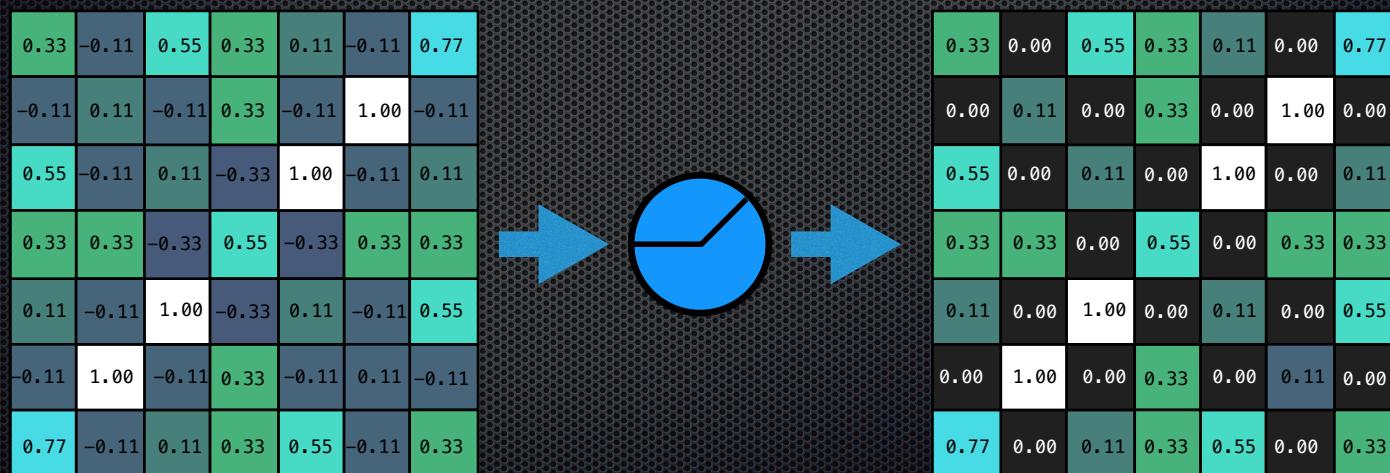
Capa ReLU:

- * Esta capa aplica la función de activación ReLU, la cual convierte todos los valores negativos a cero



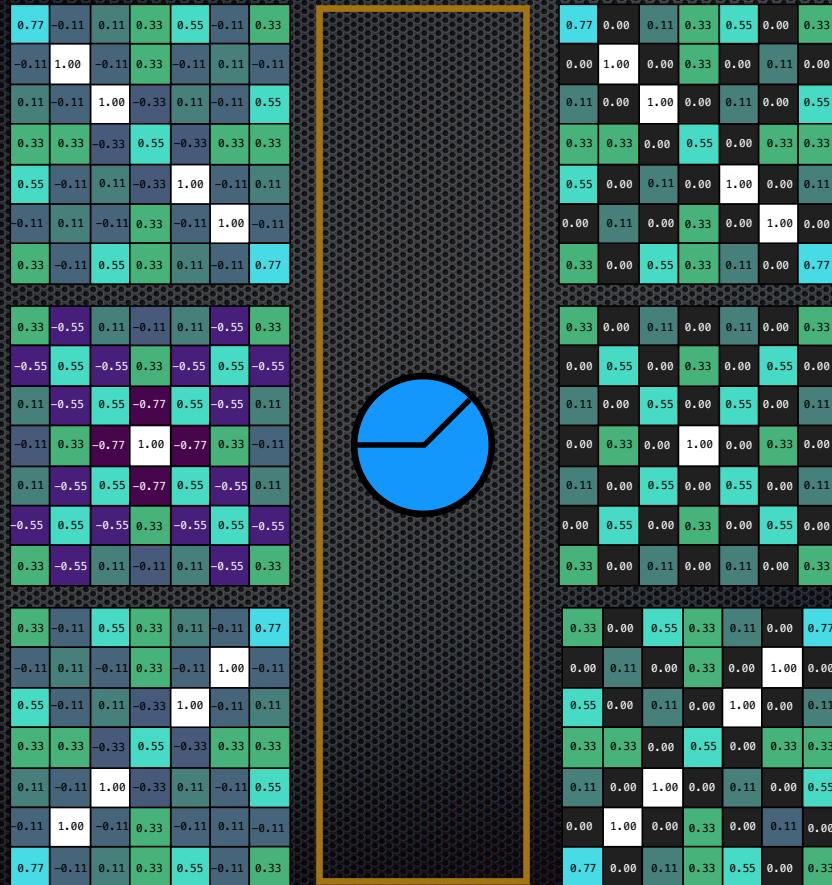
Capa ReLU:

- * Esta capa aplica la función de activación ReLU, la cual convierte todos los valores negativos a cero



Capa ReLU:

- * Una pila de imágenes se convierte en una pila de imágenes sin valores negativos



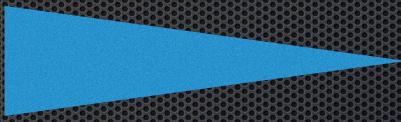
Pooling: reducción de la pila de imágenes

- * Elija un tamaño de ventana (generalmente 2×2 o 3×3).
- * Elija un paso (generalmente 2).
- * Deslice la ventana a través de sus imágenes filtradas.
- * Desde cada ventana, tome el valor máximo.

Pooling

0.77	0.00	0.11	0.33	0.55	0.00	0.33	
0.00	1.00	0.00	0.33	0.00	0.11	0.00	
0.11	0.00	1.00	0.00	0.11	0.00	0.55	
0.33	0.33	0.00	0.55	0.00	0.33	0.33	
0.55	0.00	0.11	0.00	1.00	0.00	0.11	
0.00	0.11	0.00	0.33	0.00	1.00	0.00	
0.33	0.00	0.55	0.33	0.11	0.00	0.77	

Pooling con
valor máximo



1.00	0.33	0.55	0.33
0.33	1.00	0.33	0.55
0.55	0.33	1.00	0.11
0.33	0.55	0.11	0.77

Pooling



Pooling con
valor máximo

0.55	0.33	0.55	0.33
0.33	1.00	0.55	0.11
0.55	0.55	0.55	0.11
0.33	0.11	0.11	0.33

Pooling

0.33	0.00	0.55	0.33	0.11	0.00	0.77	
0.00	0.11	0.00	0.33	0.00	1.00	0.00	
0.55	0.00	0.11	0.00	1.00	0.00	0.11	
0.33	0.33	0.00	0.55	0.00	0.33	0.33	
0.11	0.00	1.00	0.00	0.11	0.00	0.55	
0.00	1.00	0.00	0.33	0.00	0.11	0.00	
0.77	0.00	0.11	0.33	0.55	0.00	0.33	

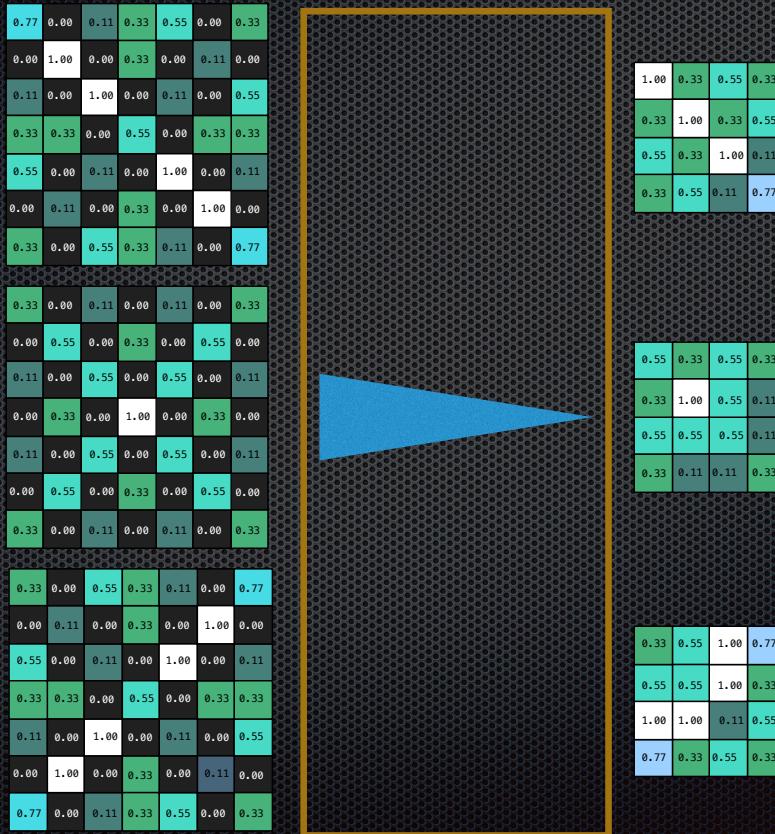
Pooling con
valor máximo



0.33	0.55	1.00	0.77
0.55	0.55	1.00	0.33
1.00	1.00	0.11	0.55
0.77	0.33	0.55	0.33

Capa de Pooling

- * Una pila de imágenes se convierte en una pila de imágenes más pequeñas



Las Capas se Apilan

- * La salida de una capa se convierte en la entrada de la próxima

-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1	
-1	-1	1	-1	-1	-1	1	-1	-1	
-1	-1	-1	1	-1	1	-1	-1	-1	
-1	-1	-1	-1	1	-1	-1	-1	-1	
-1	-1	-1	1	-1	1	-1	-1	-1	
-1	-1	1	-1	-1	-1	1	-1	-1	
-1	1	-1	-1	-1	-1	-1	1	-1	
-1	-1	-1	-1	-1	-1	-1	-1	-1	



1.00	0.33	0.55	0.33
0.33	1.00	0.33	0.55
0.55	0.33	1.00	0.11
0.33	0.55	0.11	0.77

0.55	0.33	0.55	0.33
0.33	1.00	0.55	0.11
0.55	0.55	0.55	0.11
0.33	0.11	0.11	0.33

0.33	0.55	1.00	0.77
0.55	0.55	1.00	0.33
1.00	1.00	0.11	0.55
0.77	0.33	0.55	0.33

Apilamiento Profundo

- * Las capas se pueden repetir multarles veces.

-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1



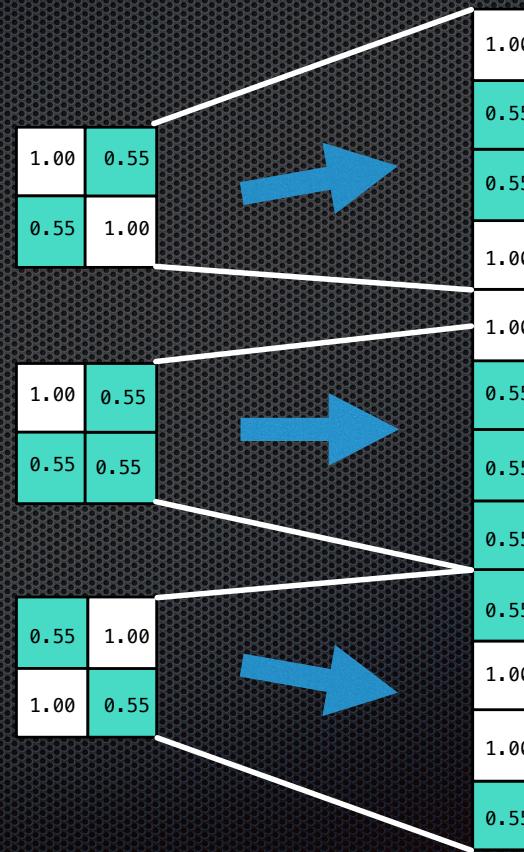
1.00	0.55
0.55	1.00

1.00	0.55
0.55	0.55

0.55	1.00
1.00	0.55

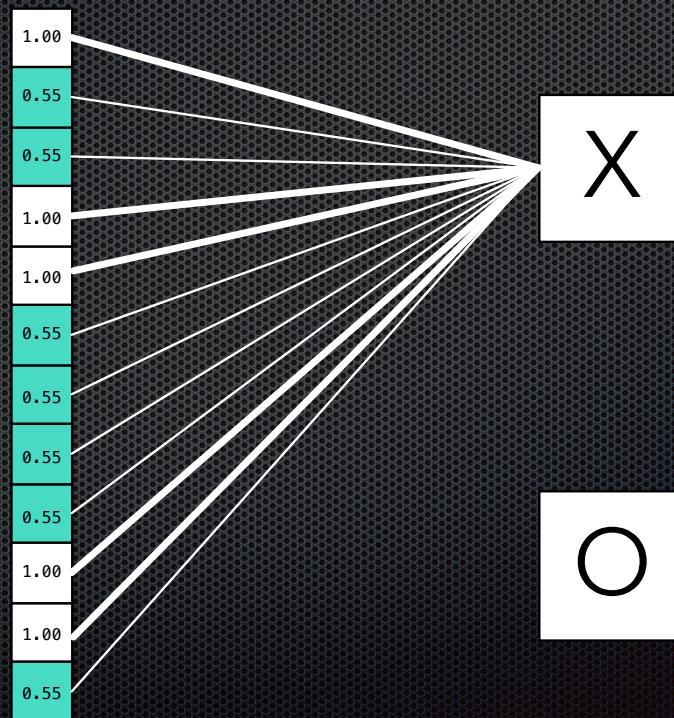
Capa Completamente Conectada

- * Cada valor tiene un voto



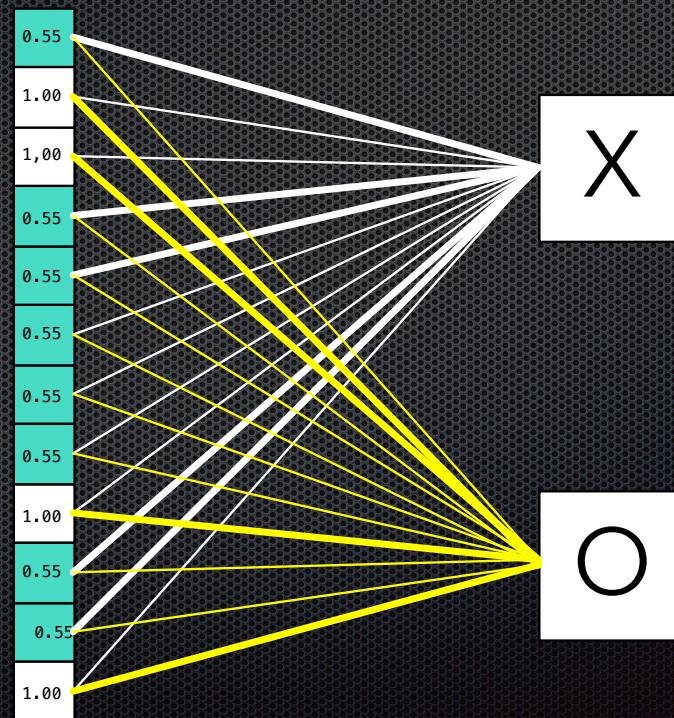
Capa Completamente Conectada

- * El voto depende de cuán fuertemente predice un valor X o O



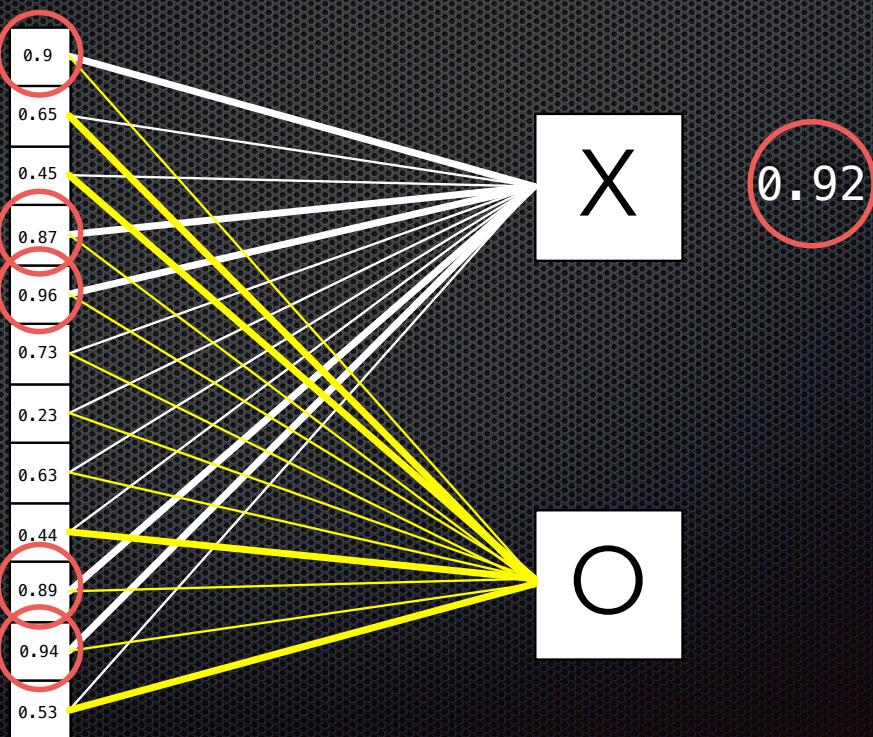
Capa Completamente Conectada

- * El voto depende de cuán fuertemente predice un valor X o O



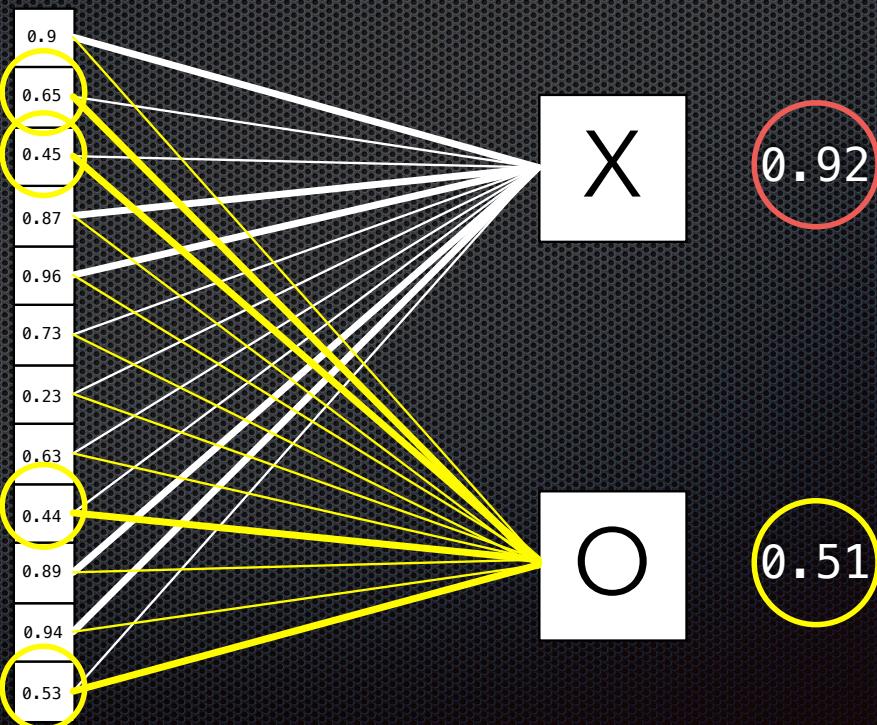
Capa Completamente Conectada

- * Valores futuros de los votes para X o O



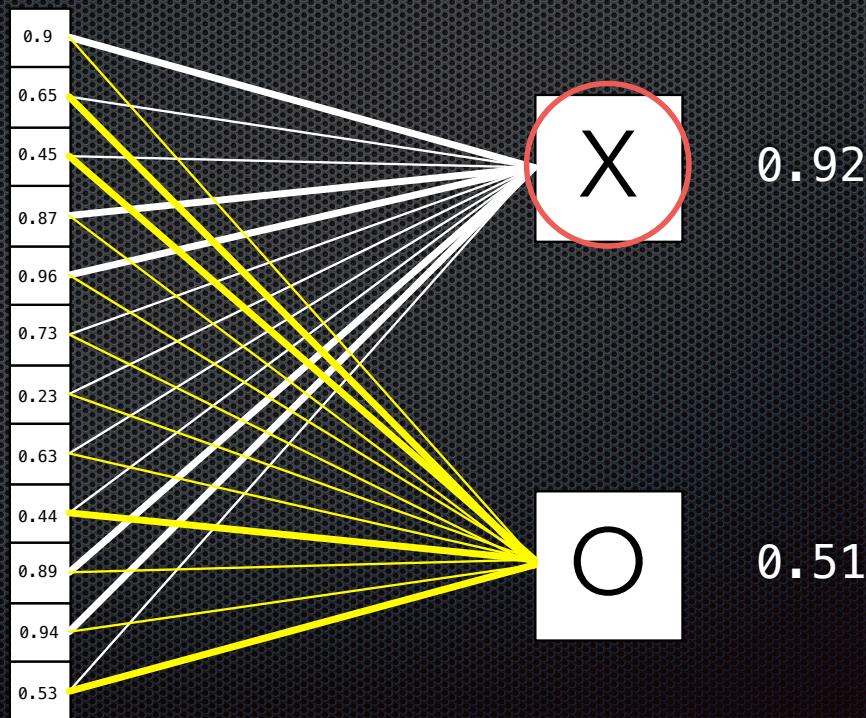
Capa Completamente Conectada

- * Valores futuros de los votes para X o O



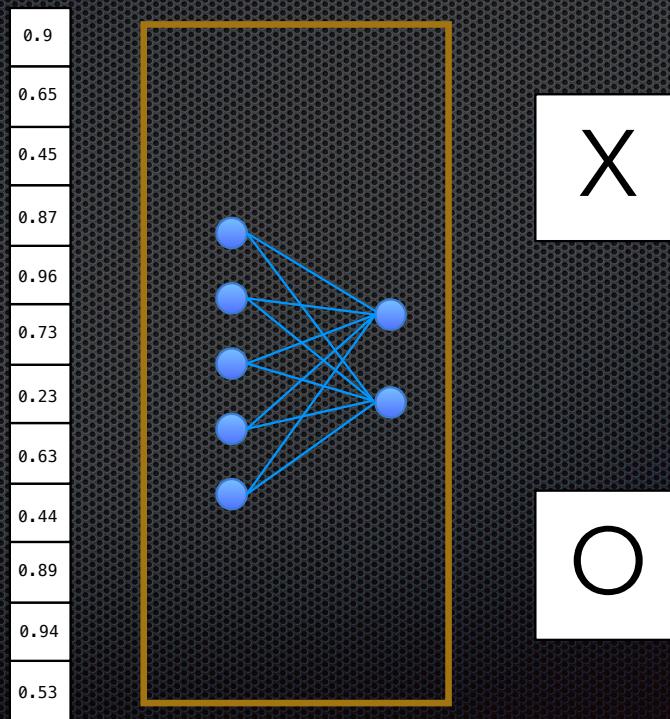
Capa Completamente Conectada

- * Valores futuros de los votes para X o O



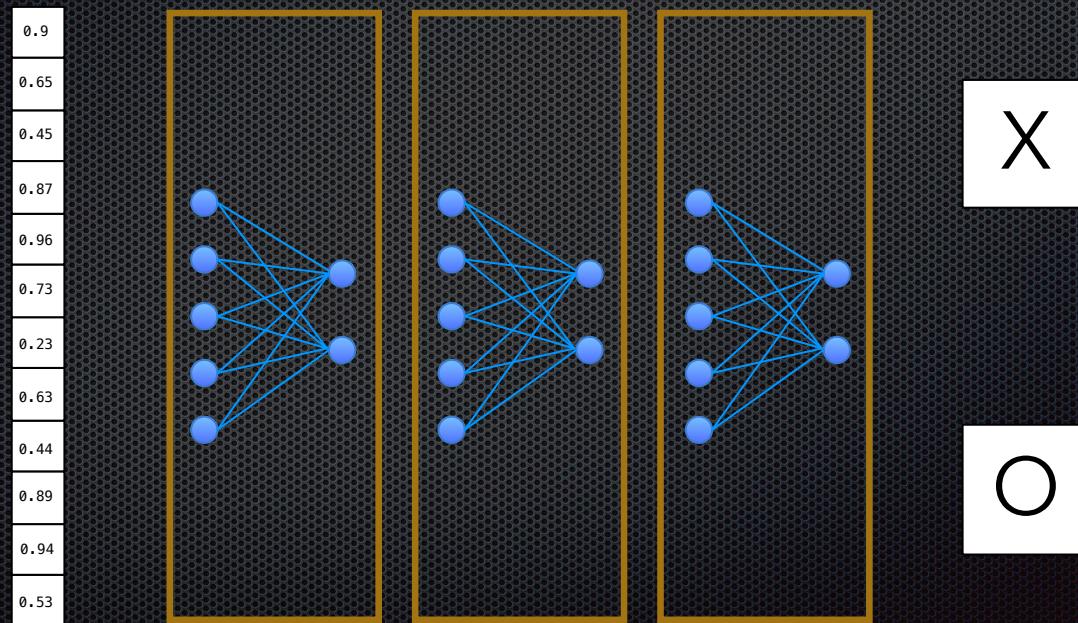
Capa Completamente Conectada

- * La lista de atributos se convierten en una lista de votos



Capa Completamente Conectada

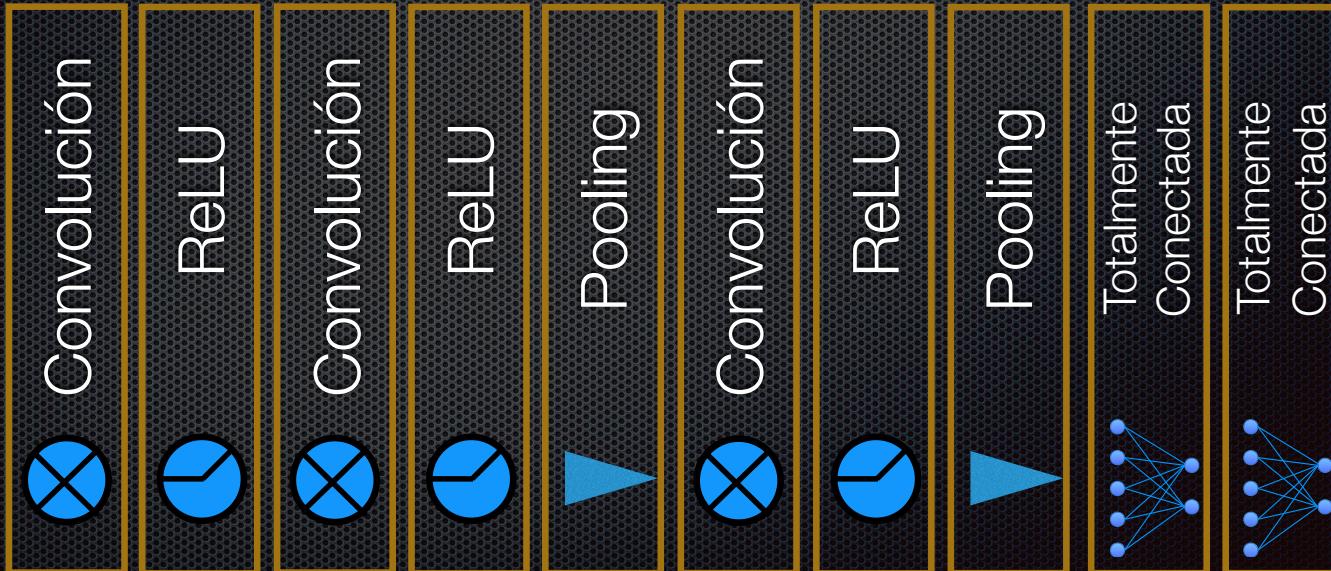
- * Esta capas también se pueden apilar



La Red Completa

- * Un conjunto de pixeles se convierten en un conjunto de votos

-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1	-1
-1	-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	1	-1	-1



0.92



0.51

Hiperparámetros

- * Convolución:
 - * Número de filtros
 - * Tamaño de los filtros
- * Pooling:
 - * Tamaño de la ventana
 - * Paso de la ventana
- * Completamente Conectada:
 - * Número de nodos

Arquitectura

- * ¿Cuantas capas de cada tipo?
- * ¿En qué orden?

