



Ayudantía 13

Redes Neuronales convolucionales

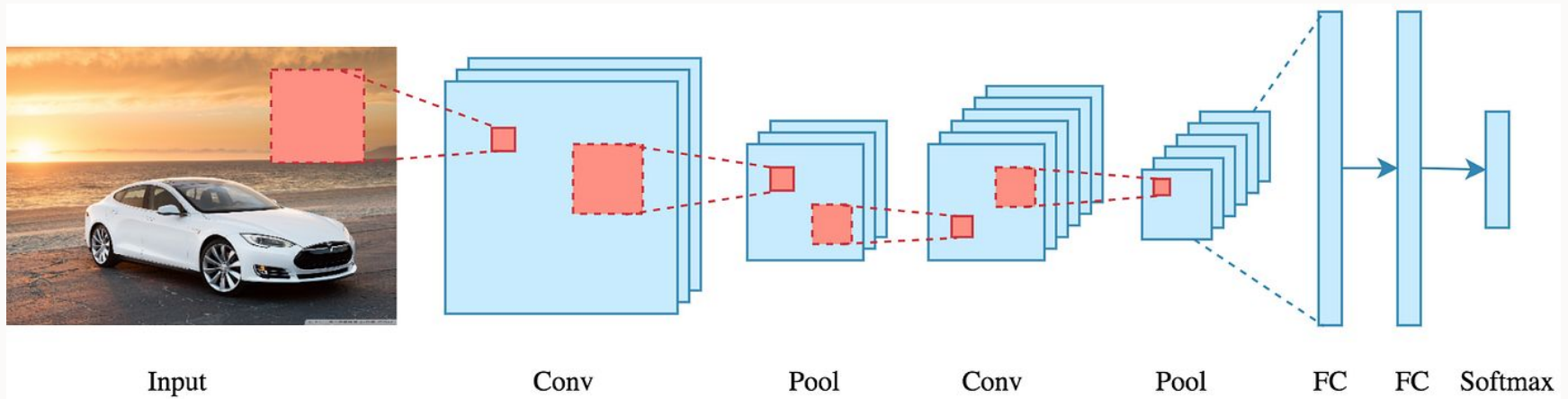
Por Bryan Acevedo y Octavio Aguila

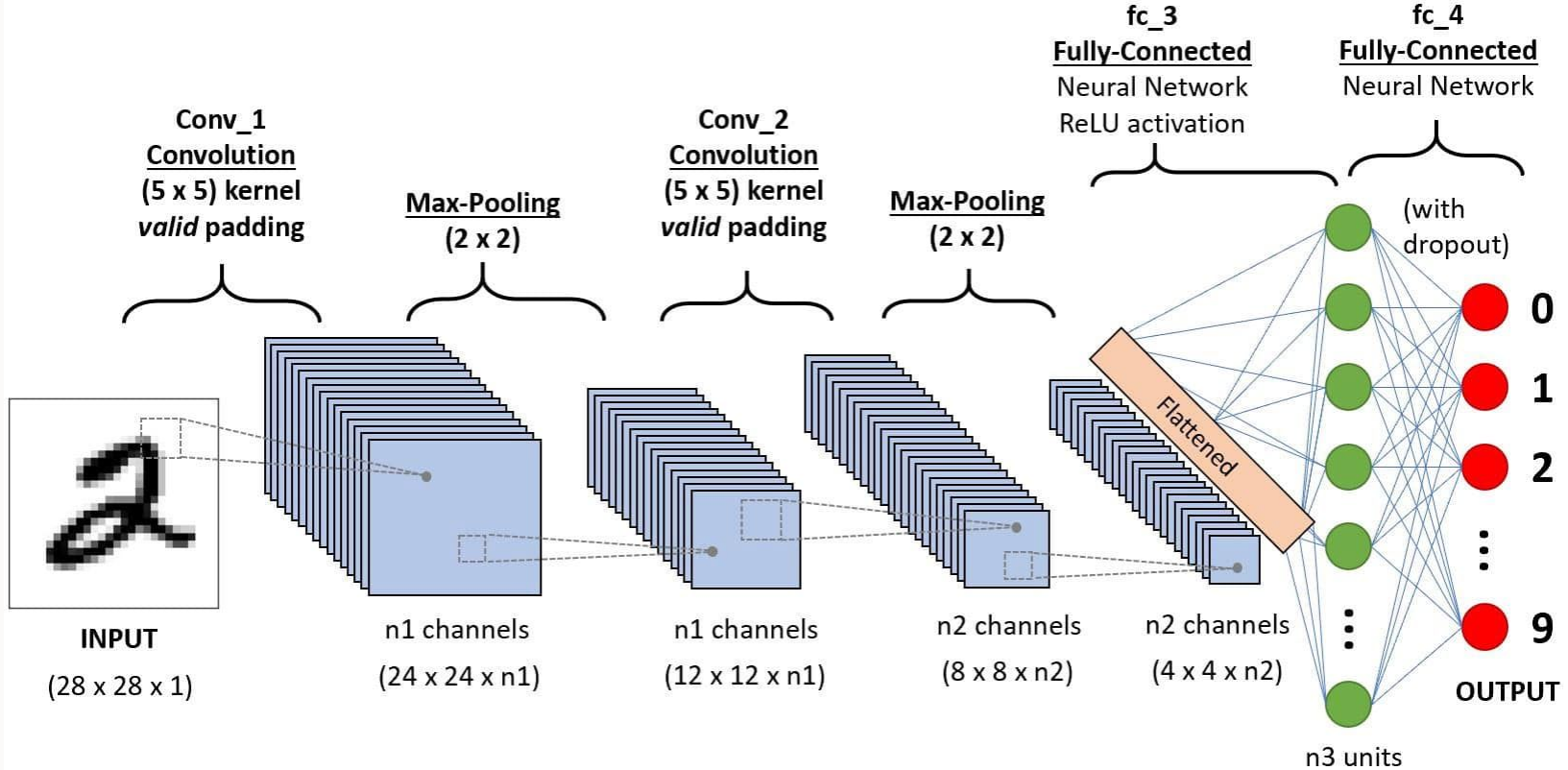
1 de diciembre 2023



¿Qué son?

Tipo de arquitectura de red neuronal especialmente diseñada para el procesamiento de datos estructurados en cuadrícula, como imágenes o audio.



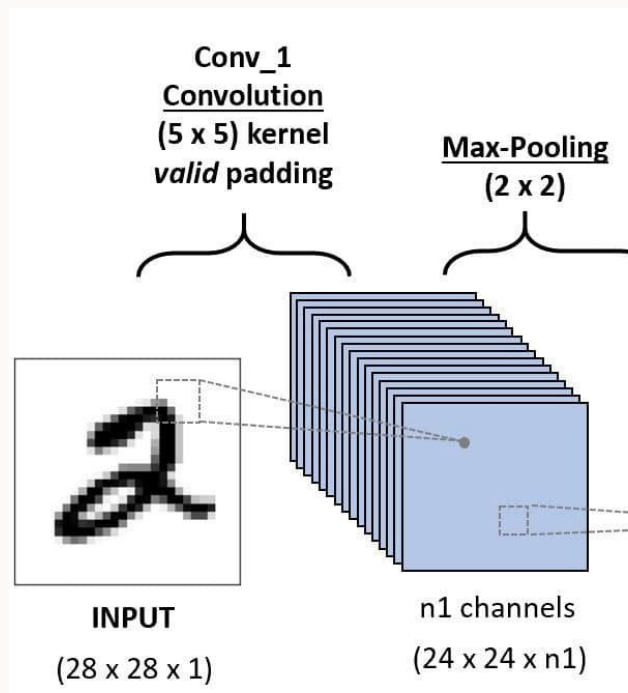
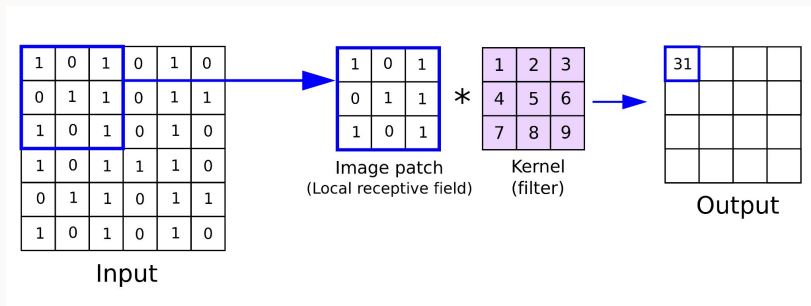




Capa Convolucional

En lugar de conectar cada neurona con todas las neuronas en la capa siguiente, las capas convolucionales utilizan filtros (kernels) que se aplican a regiones locales de la entrada.

Estos filtros se deslizan a través de la entrada para extraer características importantes.

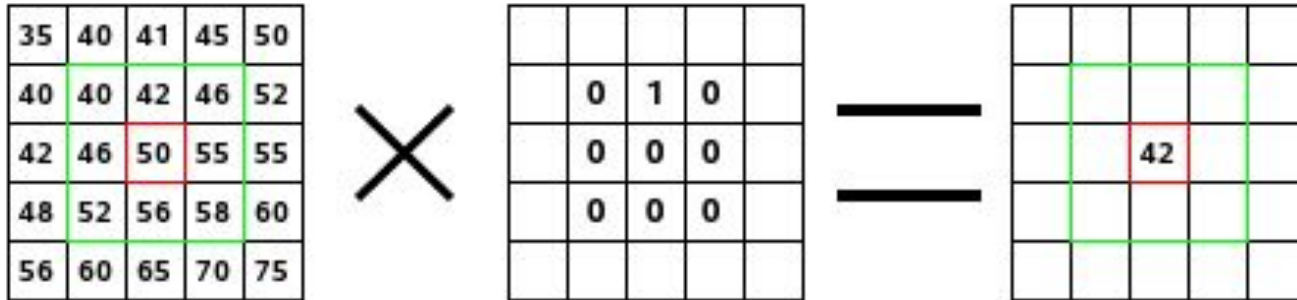




Kernels

Matrices de pesos usadas para realizar la operación de convolución, esta se desliza a través de la entrada (imagen u otra representación espacial) para realizar operaciones locales.

Cada elemento en el kernel contribuye al cálculo de un valor en la capa de salida mediante la multiplicación con los valores correspondientes en la región de la entrada que está siendo analizada, y luego sumando los resultados.





35	40	41	45	50
40	40	42	46	52
42	46	50	55	55
48	52	56	58	60
56	60	65	70	75

*

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

=

	78			

$$35 \cdot (-2) + 40 \cdot (-1) + 41 \cdot 0 + 40 \cdot (-1) + 40 \cdot 1 + 42 \cdot 1 + 42 \cdot 0 + 46 \cdot 1 + 50 \cdot 2 = 78$$

35	40	41	45	50
40	40	42	46	52
42	46	50	55	55
48	52	56	58	60
56	60	65	70	75

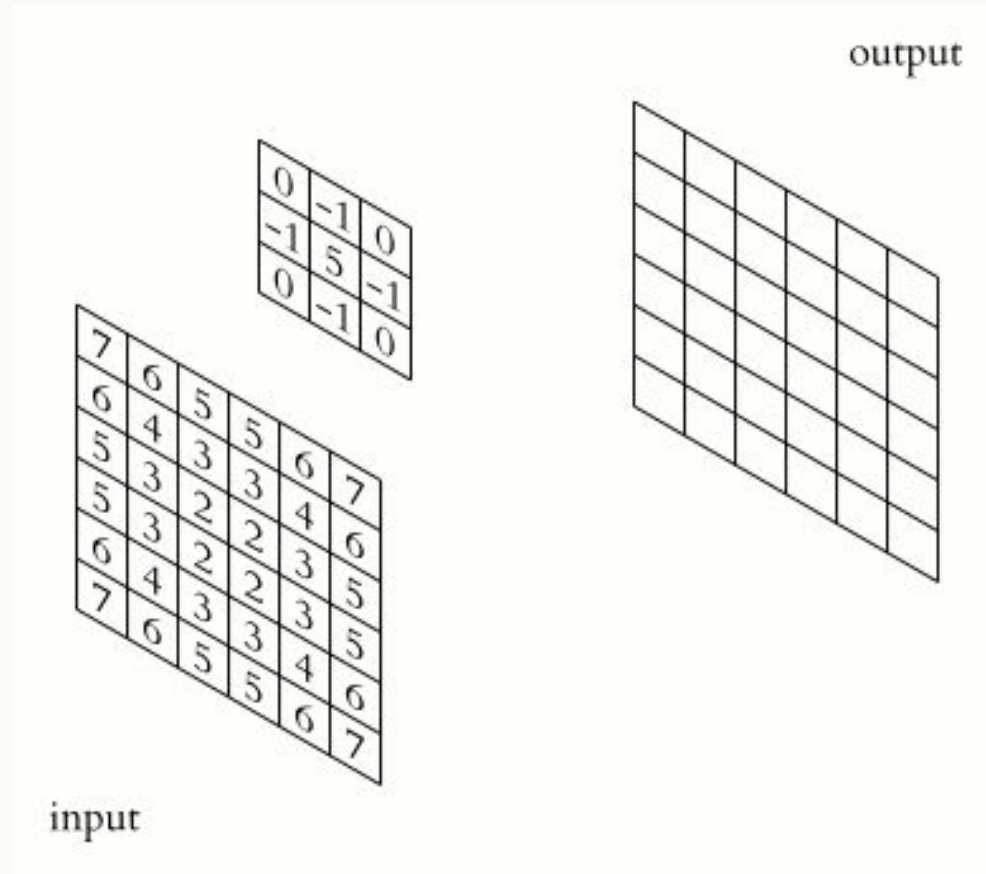
*

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

=

		87		

$$40 \cdot (-2) + 41 \cdot (-1) + 45 \cdot 0 + 40 \cdot (-1) + 42 \cdot 1 + 46 \cdot 1 + 46 \cdot 0 + 50 \cdot 1 + 55 \cdot 2 = 87$$



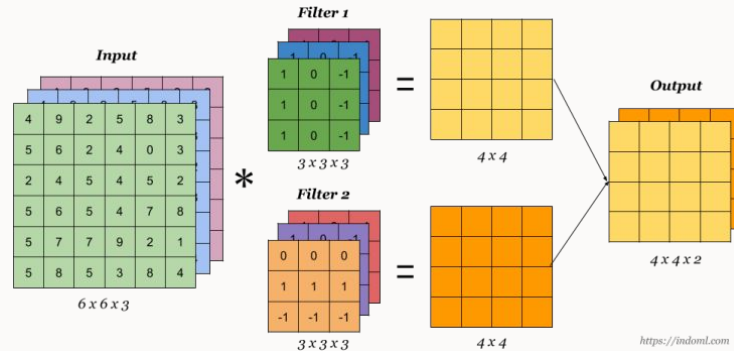
https://miro.medium.com/v2/resize:fit:780/1*Eai425FYQQSNOaahTXtgq.gif



Operación Convolución (Explicación)

Consideremos una imagen cuyas dimensiones son alto, ancho y profundidad, el detector de características podrá ser una matriz de 2 dimensiones. El kernel se aplica sobre un área de los datos y se calcula el producto punto entre estos datos seleccionados y el filtro.

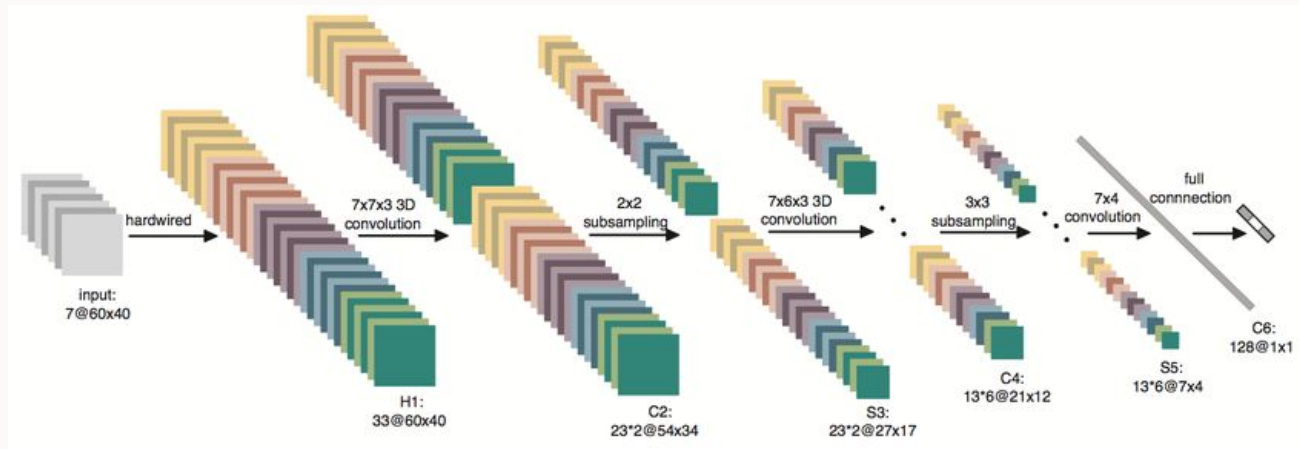
El resultado será un escalar que se introduce en una matriz de salida. Luego el filtro cambia los datos a los que va a realizar la operación, repitiendo el proceso hasta que se ha recorrido todos los datos. Esta es la matriz final.





Operación Convolución (Explicación)

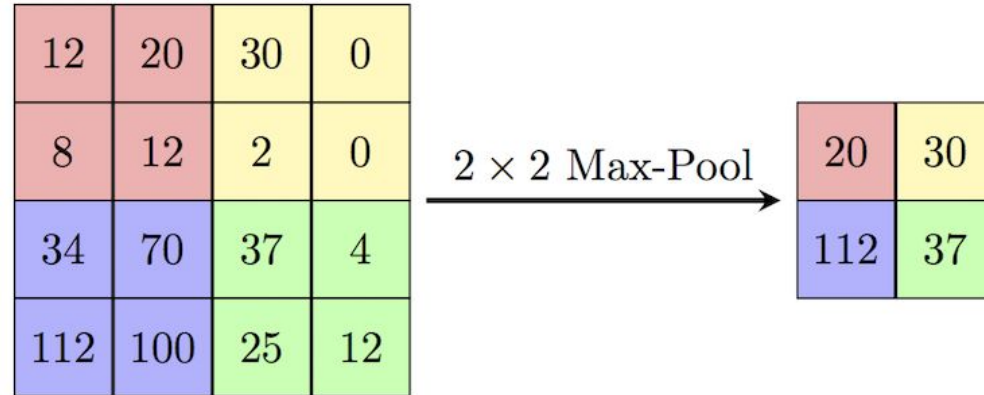
Ahora, considerando los pesos en el detector de características, se tiene que estos permanecen fijos a medida que se mueven por los datos de entrada (uso compartido de parámetros) y estos se pueden ajustar durante el entrenamiento utilizando metodologías, como por ejemplo, backpropagation y descenso del gradiente.





Pooling

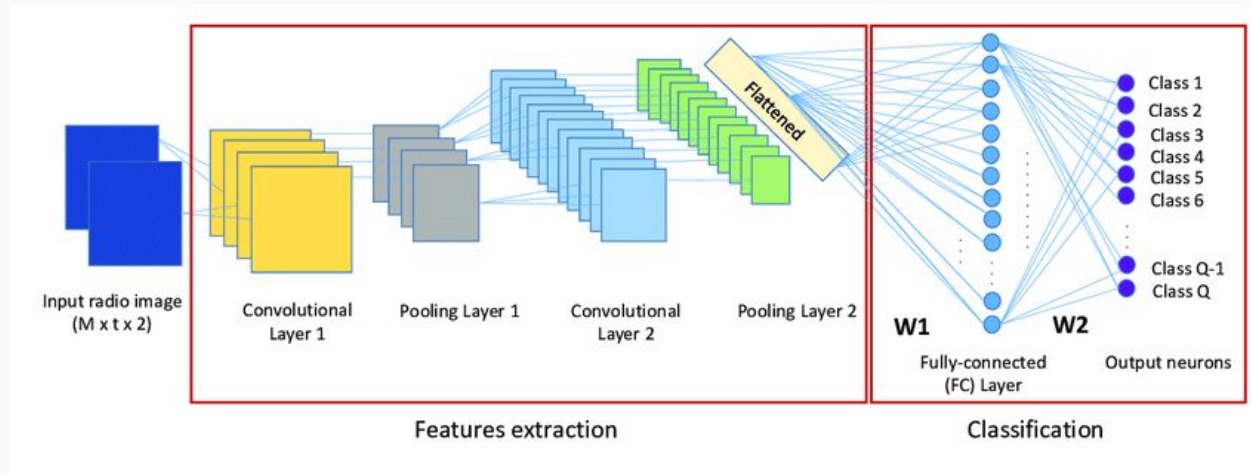
Después de las capas convolucionales, generalmente se aplica una capa de agrupación para reducir la dimensionalidad y la cantidad de parámetros en la red. La operación de agrupación (pooling) reduce el tamaño espacial de la representación y, al mismo tiempo, mantiene las características más importantes.





Capa completamente conectada

Utilización de una capa de redes neurales completamente conectadas, las cuales utilizan ReLU (Rectified Linear Unit), para introducir no linealidades en la red y permitirle aprender patrones más complejos.





Funciones de pérdida

Se utiliza la entropía cruzada, para medir la diferencia entre las predicciones y las etiquetas reales. Para obtener una distribución de probabilidad de nuestras predicciones utilizamos la función softmax.

$$H(p, q) = - \sum_{x \in \text{classes}} p(x) \log q(x)$$

True probability distribution (one-shot)

Your model's predicted probability distribution



Entrenamiento

Las CNN se entrenan mediante el ajuste de los pesos de los filtros y las conexiones entre las neuronas utilizando algoritmos de optimización, como el descenso del gradiente estocástico (SGD) o variantes más avanzadas como Adam.

