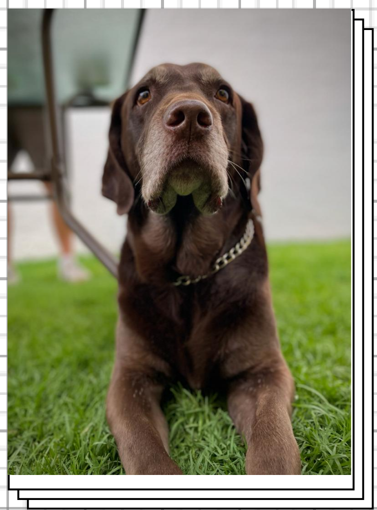
Redes convolucionales 2D [CNNs 2D]

- Descripción de Imágenes digitalizadas.
- Propiedades del sistema RGB (*Red-Green-Blue*).



$$\begin{bmatrix} I_{1,1} & I_{1,2} & \cdots & I_{1,h} \\ I_{2,1} & I_{2,2} & \cdots & I_{2,h} \\ \vdots & \vdots & & \vdots \\ I_{w,1} & I_{w,2} & \cdots & I_{w,h} \end{bmatrix} \begin{bmatrix} R_{1,1} & R_{1,2} & \cdots & R_{1,h} \\ R_{2,1} & R_{2,2} & \cdots & R_{2,h} \\ \vdots & \vdots & & \vdots \\ R_{w,1} & R_{w,2} & \cdots & R_{w,h} \end{bmatrix} \begin{bmatrix} G_{1,1} & G_{1,2} & \cdots & G_{1,h} \\ G_{2,1} & G_{2,2} & \cdots & G_{2,h} \\ \vdots & \vdots & & \vdots \\ G_{w,1} & G_{w,2} & \cdots & G_{w,h} \end{bmatrix} \begin{bmatrix} B_{1,1} & B_{1,2} & \cdots & B_{1,h} \\ B_{2,1} & B_{2,2} & \cdots & B_{2,h} \\ \vdots & \vdots & & \vdots \\ B_{w,1} & B_{w,2} & \cdots & B_{w,h} \end{bmatrix}$$

$$I_{a,b} = [R_{a,b}, G_{a,b}, B_{a,b}] \in [0,255] \times [0,255] \times [0,255]$$

$$I \in \mathbb{R}^{h \times w \times c} \qquad (h = w = n, c = 3)$$

Capas convolucionales (para un solo canal)

- Invarianza de traslación (*translation invariance*).
- Principio de localidad (*locality principle*).

Sea $X = (x_{k,l}) \in \mathbb{R}^{n\times n\times 1}$ una imagen de $n \times n$ pixeles y $Z = (z_{i,j}) \in \mathbb{R}^{m\times m\times 1}$ la variable resultante de aplicarle una capa convolucional a la imagen. Establecemos $m < n \in \mathbb{Z}^{+}$.

$$z_{i,j} = u_{i,j} + \sum_{k=1}^{n}\sum_{l=1}^{n} w_{i,j,k,l} \cdot x_{k,l}$$

$$= u_{i,j} + \sum_{a=1}^{n}\sum_{b=1}^{n} w_{i,j,a,b} \cdot x_{i+a,j+b} \quad k = i + a \quad l = j + b$$

$$= u + \sum_{a=1}^{n}\sum_{b=1}^{n} w_{a,b} \cdot x_{i+a,j+b}$$

$$= u + \sum_{a=1}^{n}\sum_{b=1}^{n} w_{a,b} \cdot 1_{A} \cdot x_{i+a,j+b} \quad A = (|a| > \Delta_i) \cup (|b| > \Delta_j)$$

Ejemplo

$$\underset{n=3}{\overset{X}{\begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{bmatrix}}} * \overset{\text{kernel}}{\begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}} = \underset{m=2}{\overset{Z}{\begin{bmatrix} & \\ & \end{bmatrix}}}$$

Operador de convolución

$$(f * g)(\mathbf{x}) = \int_{\mathbb{R}^d} f(\mathbf{y})g(\mathbf{x} - \mathbf{y})dy$$

$$(f * g)(i, j) = \sum_a \sum_b f(a, b)g(i - a, j - b)$$

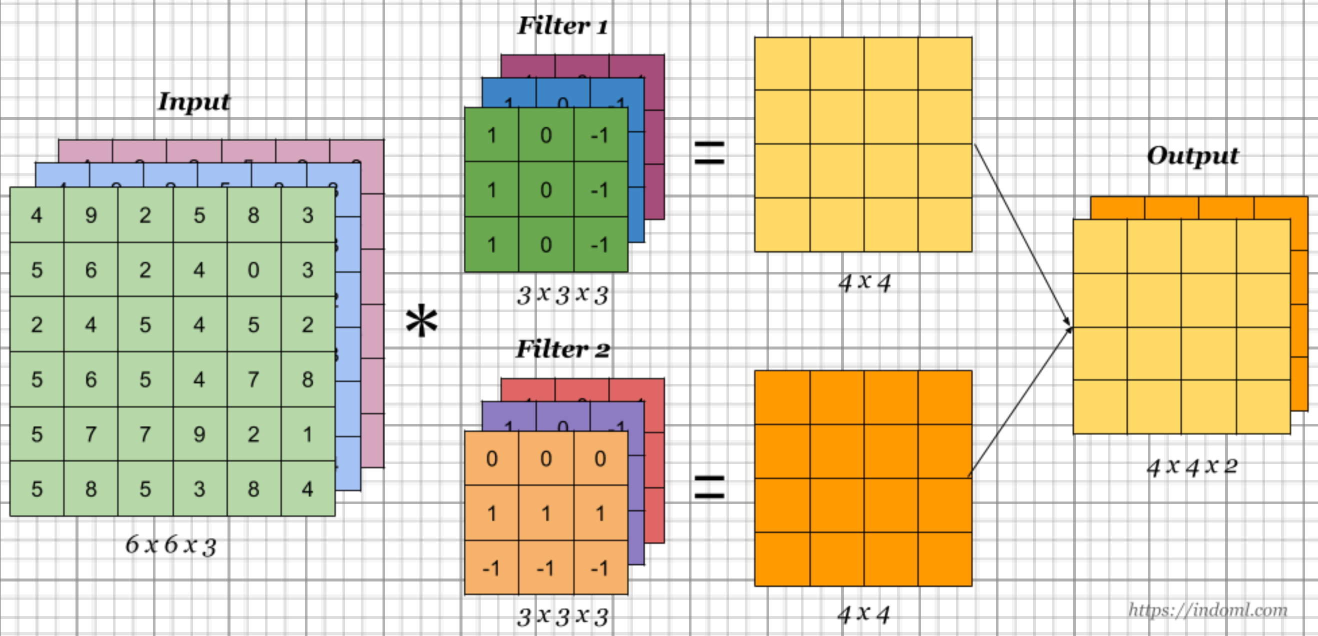
Capas convolucionales (múiltiples canales).

$$z_{i,j,d} = u + \sum_{a \in A} \sum_{b \in A} \sum_{c} w_{a,b,c,d} \cdot x_{i+a,j+b,c}$$

$$n \times n \times c \quad * \quad (k \times k \times c) \times d \quad \to \quad (n-k+1) \times (n-k+1) \times d$$

$$w \times h \times c \quad * \quad (k_w \times k_h \times c) \times d \quad \to \quad (w-k_w+1) \times (h-k_h+1) \times d$$

Ejemplo


Input
4 9 2 5 8 3
5 6 2 4 0 3
2 4 5 4 5 2
5 6 5 4 7 8
5 7 7 9 2 1
5 8 5 3 8 4
6 x 6 x 3
*
Filter 1
1 0 -1
1 0 -1
1 0 -1
3 x 3 x 3
Filter 2
0 0 0
1 1 1
-1 -1 -1
3 x 3 x 3
=
4 x 4
=
4 x 4
Output
4 x 4 x 2
https://indoml.com