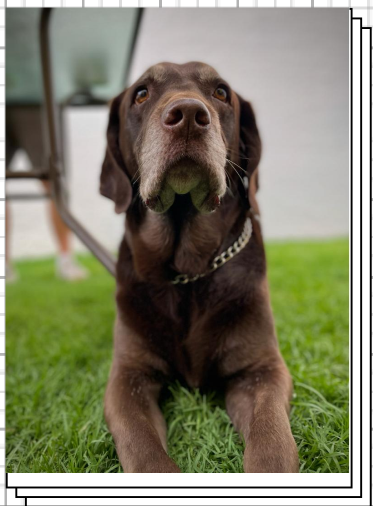
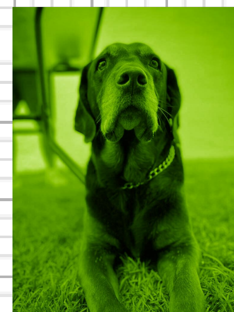
Redes convolucionales 2D [CNNs 2D]

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

$$\frac{R+G+B}{3} \to I$$



I_{C_R}



I_{C_G}

I_{C_B}

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

$2^8 - 1$

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

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

Capas convolucionales (para un solo canal)

(1). Invarianza de traslación (*translation invariance*).
(2). 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 $m < n \in \mathbb{Z}^+$.

$$Z = Wx + b$$

$$z_{i,j} = u_{i,j} + \sum_{k=1}^{n}\sum_{l=1}^{n} w_{i,j,k,l}\cdot x_{k,l} \longrightarrow \text{Feed Forward Fully connected}$$

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

constante para cada (i,j)

$$= u + \sum_{a=1}^{n}\sum_{b=1}^{n} w_{a,b}\cdot x_{i+a,j+b} \longrightarrow (1) \qquad \Delta_i = ?$$

$$\longrightarrow (2) \qquad i \qquad j$$

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

Kernel
Filtro (Filter)

$$|a| > \Delta \longrightarrow a > \Delta \text{ o } a < -\Delta$$

Ejemplo

$$\underset{n=3}{\overset{X}{\begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{bmatrix}}} * \overset{\text{kernel}}{\begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}} = \underset{m=2}{\overset{Z}{\begin{bmatrix} 19 & 25 \\ 33 & 43 \end{bmatrix}}}$$

$z_{1,1} = 0 \cdot 0 + 1 \cdot 1 + 3 \cdot 2 + 4 \cdot 3$
$= 19$

$z_{2,1} = 3 \cdot 0 + 4 \cdot 1 + 6 \cdot 2 + 7 \cdot 3$
$= 33$

$z_{1,2} = 0 + 2 + 8 + 15$
$= 25$

$z_{22} = 0 + 5 + 14 + 21$
$= 43$

Operador de convolución

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

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

Capas convolucionales (múiltiples canales).

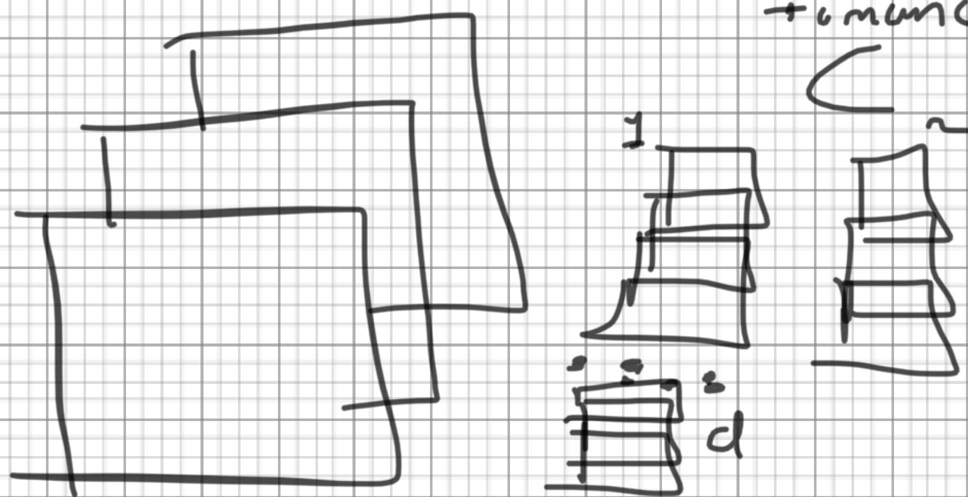
$$z_{i,j,d} = u + \sum_{a\in A}\sum_{b\in A}\sum_{c} w_{a,b,c,d} \cdot x_{i+a,j+b,c}$$

$X = (x_{k,l,c}) \in \mathbb{R}^{n\times n\times c}$

C c

$Z = (z_{i,j,d}) \in \mathbb{R}^{m\times m\times D}$

$D \to$ kernel $\downarrow$ tamaño C



$D :=$ # canales de (Z) la salida

$C :=$ # canales de (X) entrada

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

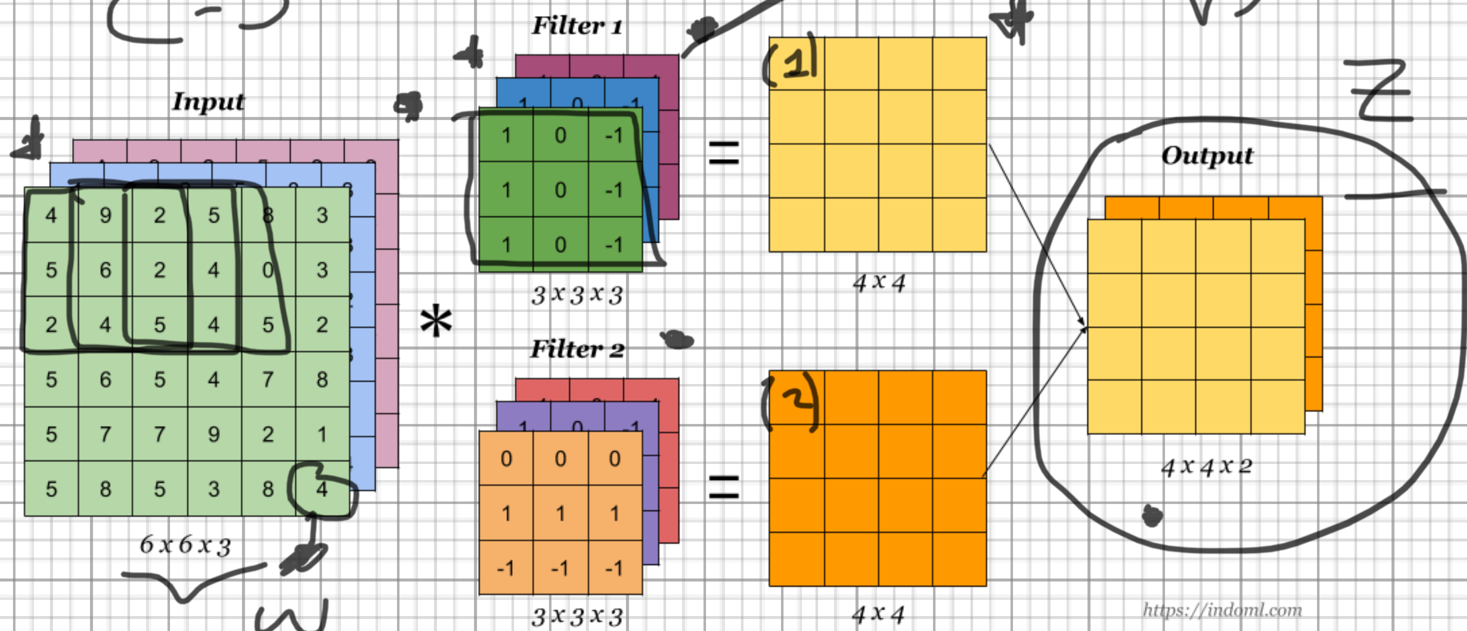
$$w \times h \times c \quad * \quad (k_w \times k_h \times c) \times d \quad \to \quad (w-k_w+1)\times(h-k_h+1)\times d$$

Ejemplo

$C=3$

$D=1$

$\left.\begin{array}{l} M \\ A \\ V \end{array}\right\} + (1)$



Z

W

$k=3$

$(6-3+1)=4$

Z_{ij} depender alrededor de $X(ij)$