

# APRENDIZAJE PROFUNDO

## VARIANTES DE CAPAS CONVOLUCIONALES

---

Gibran Fuentes-Pineda

Septiembre 2025

- Filtros que pueden descomponerse como el producto de 2 más simples. Por ej., un filtro 2D puede separarse en 2 filtros 1D.

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \times \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$$

# CONVOLUCIÓN ESPACIAL SEPARABLE

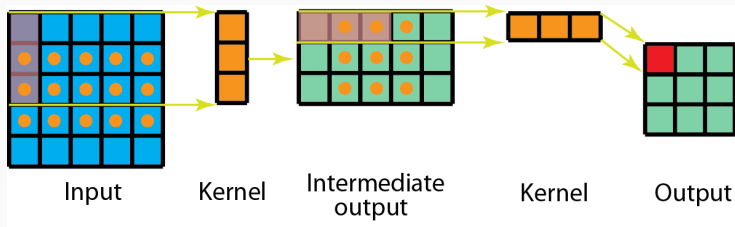


Imagen tomada de Kunlun Bai. A Comprehensive Introduction to Different Types of Convolutions in Deep Learning, 2019.

# CONVOLUCIÓN EN PROFUNDIDAD

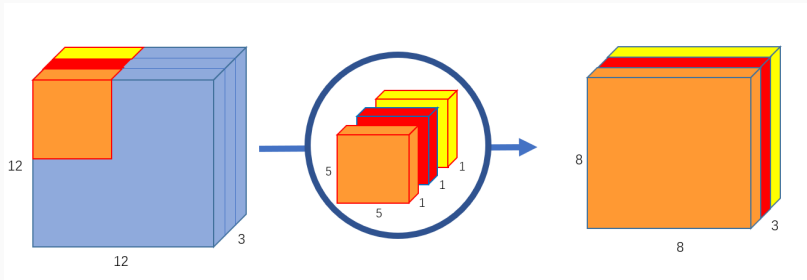


Imagen tomada de Chi-Feng Wang, A Basic Introduction to Separable Convolutions, 2018.

# CONVOLUCIÓN PUNTUAL (DE $1 \times 1$ )

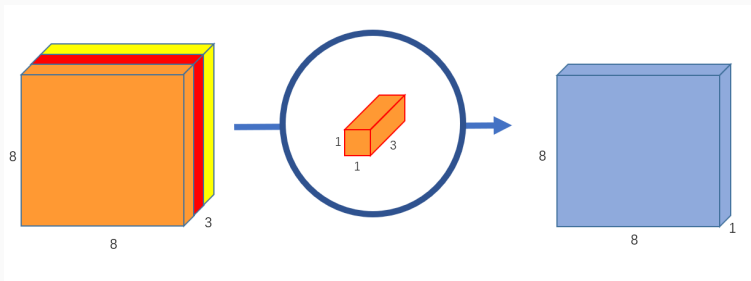


Imagen tomada de Chi-Feng Wang, A Basic Introduction to Separable Convolutions, 2018.

# CONVOLUCIÓN PUNTUAL (DE $1 \times 1$ )

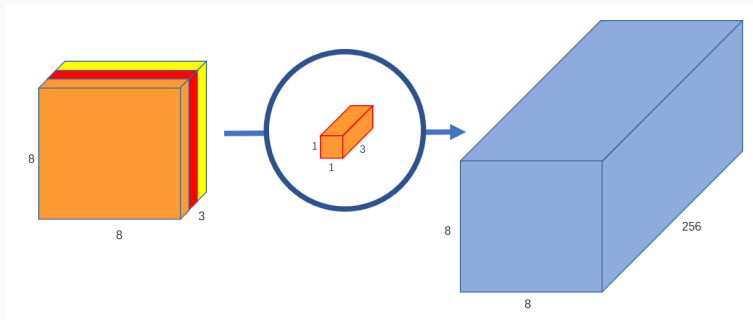


Imagen tomada de Chi-Feng Wang, A Basic Introduction to Separable Convolutions, 2018.

# CONVOLUCIÓN SEPARABLE EN PROFUNDIDAD

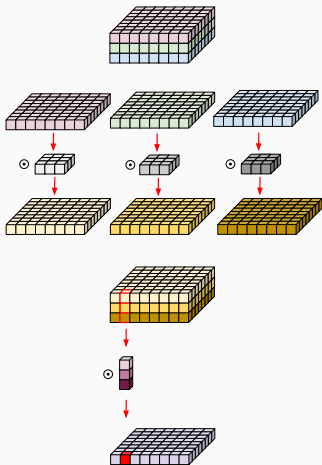
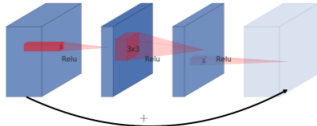


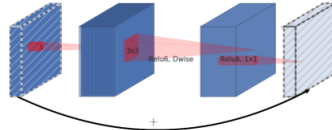
Imagen tomada de Eli Bendersky. Depthwise separable convolutions for machine learning, 2018.

# CUELLOS DE BOTELLA RESIDUALES INVERSOS

(a) Residual block



(b) Inverted residual block



Input	Operator	Output
$h \times w \times k$	1x1 conv2d, ReLU6	$h \times w \times (tk)$
$h \times w \times tk$	3x3 dwse s=s, ReLU6	$\frac{h}{s} \times \frac{w}{s} \times (tk)$
$\frac{h}{s} \times \frac{w}{s} \times tk$	linear 1x1 conv2d	$\frac{h}{s} \times \frac{w}{s} \times k'$

Imagen y tabla tomadas de Sandler et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks, 2018.



# EFFICIENTNETS (1)

- Emplea bloques de cuello de botella residuales inversos
- Explora el crecimiento de la red en tres dimensiones: profundidad ( $d$ ), amplitud ( $w$ ) y resolución de la entrada ( $r$ )

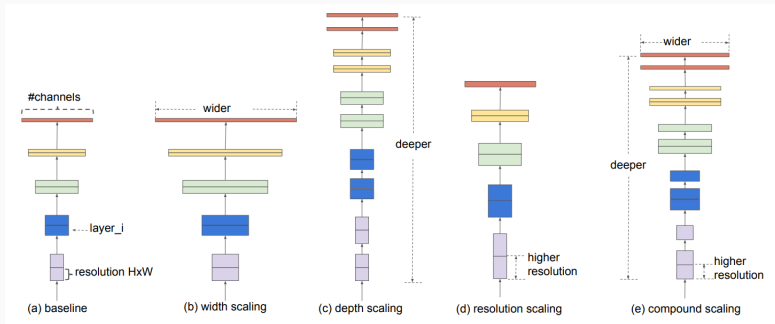


Tabla tomada de Tan and Le. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, 2019.

## EFFICIENTNETS (2)

- Exactitud para diferentes profundidades, amplitudes y resoluciones de entrada

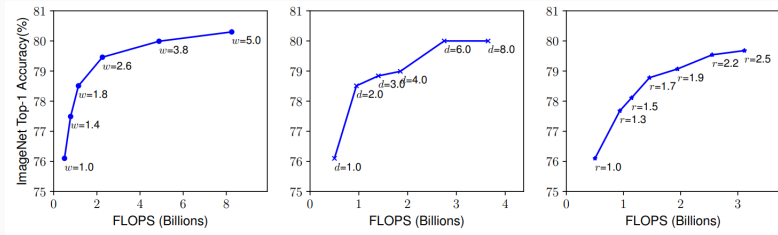


Tabla tomada de Tan and Le. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, 2019.

## EFFICIENTNETS (3)

- Búsqueda de arquitectura neuronal (NAS), optimizando de forma conjunta eficiencia y exactitud
  - Versiones B0–B7 de arquitectura EfficientNet, siendo la más básica la B0

**Table 1. EfficientNet-B0 baseline network** – Each row describes a stage  $i$  with  $\hat{L}_i$  layers, with input resolution  $(\hat{H}_i, \hat{W}_i)$  and output channels  $\hat{C}_i$ . Notations are adopted from equation 2.

Stage $i$	Operator $\hat{\mathcal{F}}_i$	Resolution $\hat{H}_i \times \hat{W}_i$	#Channels $\hat{C}_i$	#Layers $\hat{L}_i$
1	Conv3x3	$224 \times 224$	32	1
2	MBConv1, k3x3	$112 \times 112$	16	1
3	MBConv6, k3x3	$112 \times 112$	24	2
4	MBConv6, k5x5	$56 \times 56$	40	2
5	MBConv6, k3x3	$28 \times 28$	80	3
6	MBConv6, k5x5	$14 \times 14$	112	3
7	MBConv6, k5x5	$14 \times 14$	192	4
8	MBConv6, k3x3	$7 \times 7$	320	1
9	Conv1x1 & Pooling & FC	$7 \times 7$	1280	1

Tabla tomada de Tan and Le. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, 2019.

- Búsqueda de arquitectura neuronal (NAS) con bloques MBConv y Fuse-MBConv, optimizando de forma conjunta eficiencia, exactitud y velocidad de entrenamiento

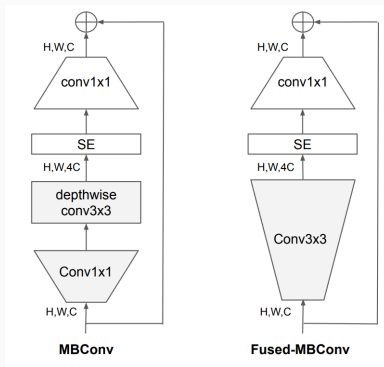


Tabla tomada de Tan and Le. EfficientNetV2: Smaller Models and Faster Training, 2021.

# BLOQUES CONVNEXT

- La familia de arquitecturas ConvNeXt actualiza la familia ResNet con bloques cuello de botella residuales inversos y otros hiperparámetros del entrenamiento

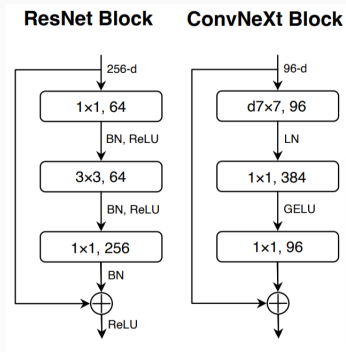


Imagen tomada de Liu et al. A ConvNet for the 2020s, 2022.

# CONVOLUCIÓN DILATADA

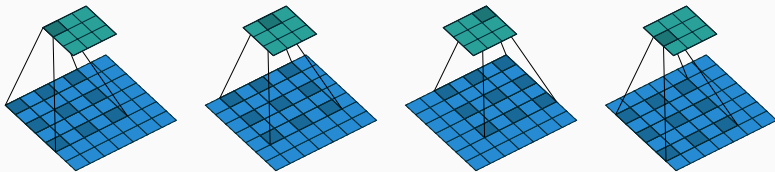


Imagen tomada de Dumoulin and Visin. A guide to convolution arithmetic for deep learning, 2018.

# CONVOLUCIÓN TRANSPUESTA (1)

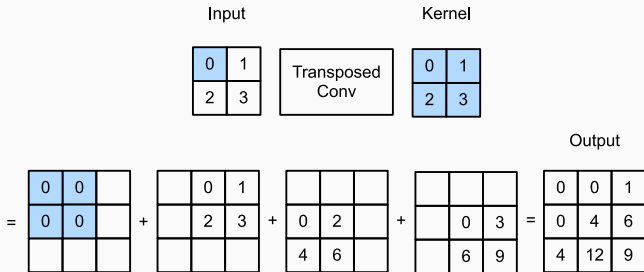


Figura tomada de Zhang et al. Dive into Deep Learning, 2022.

# CONVOLUCIÓN TRANSPUESTA (2)

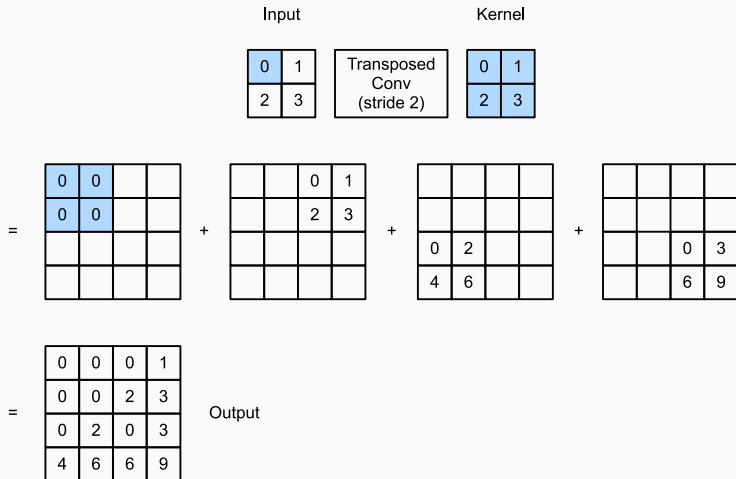


Figura tomada de Zhang et al. Dive into Deep Learning, 2022.



# REDES CONVOLUCIONALES 1D

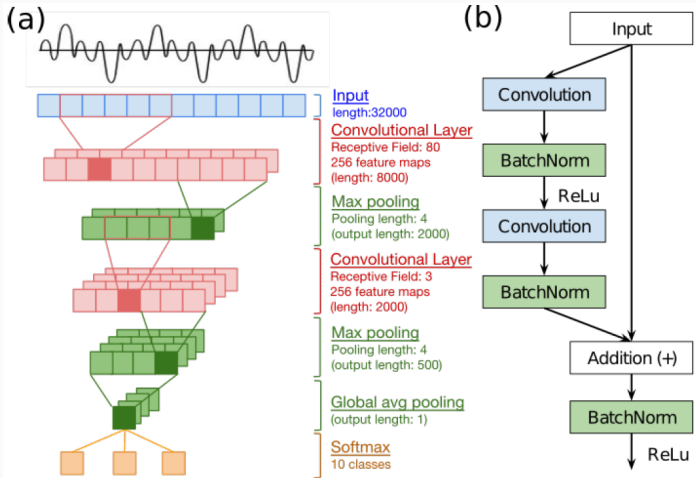


Imagen tomada de Dai et al. *Very Deep Convolutional Neural Networks for Raw Waveforms*, 2016

# REDES CONVOLUCIONALES 3D (1)

- Respuesta en redes convolucionales 2D en una imagen

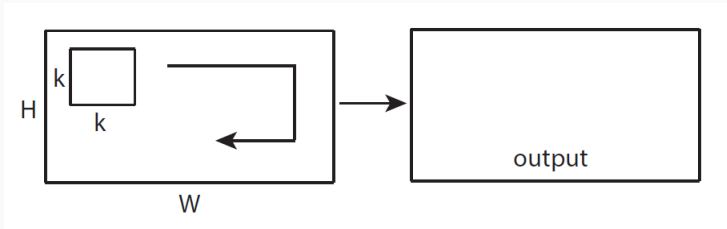


Imagen tomada de Tran et al. *Learning Spatiotemporal Features with 3D Convolutional Networks*, 2015

## REDES CONVOLUCIONALES 3D (2)

- Respuesta en redes convolucionales 2D en varias imágenes

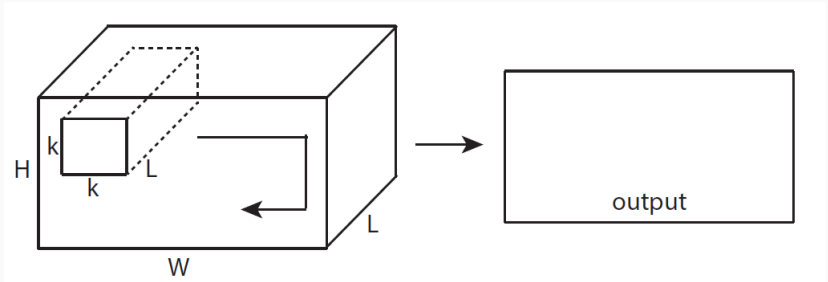


Imagen tomada de Tran et al. *Learning Spatiotemporal Features with 3D Convolutional Networks*, 2015

## REDES CONVOLUCIONALES 3D (3)

- Respuesta en redes convolucionales 3D

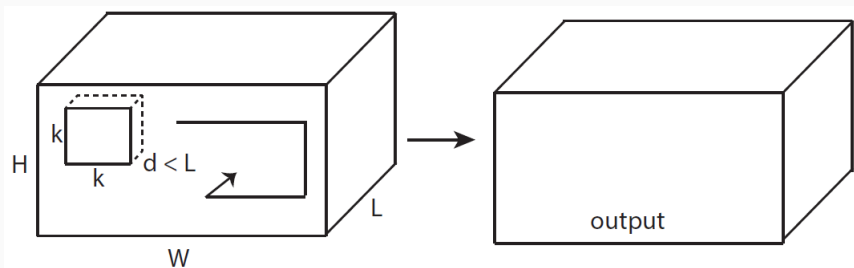


Imagen tomada de Tran et al. *Learning Spatiotemporal Features with 3D Convolutional Networks*, 2015