

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Posgrado en Ciencia e Ingeniería de la Computación

Aprendizaje Profundo

Redes convolucionales

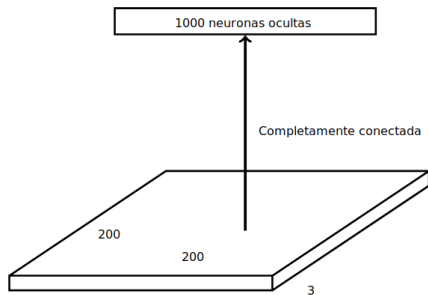
Profesor: Gibran Fuentes Pineda

Ayudantes: Berenice & Ricardo Montalvo Lezama

Septiembre 2021

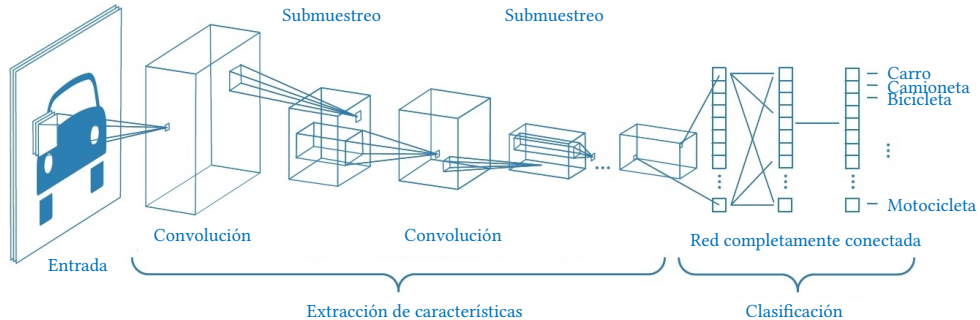
Problemática de usar redes densas para imágenes

- Supongamos que queremos entrenar una red que tome una imagen RGB de 200×200 .

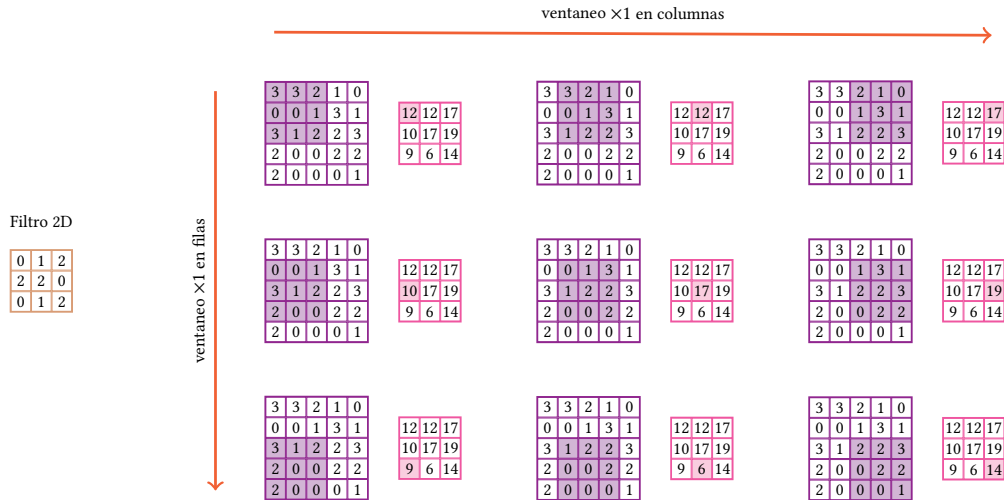


- ¡Se requieren muchos parámetros!
 - Entrada = $200 \times 200 \times 3 = 120,000$.
 - Parámetros = $120,000 \times 1000 + 1000 = 120,001,000$.

Red neuronal convolucional

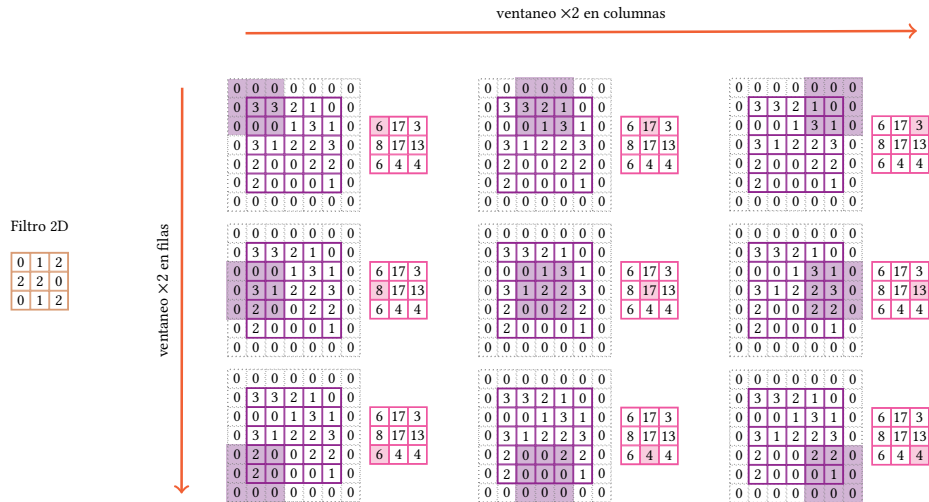


Convolución 2D



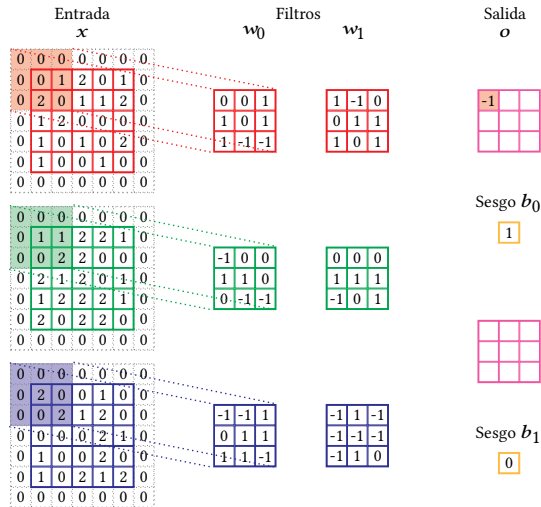
Convolución: entrada 5×5 , filtro 3×3 , paso 1×1 , salida 3×3 ,

Convolución 2D con relleno

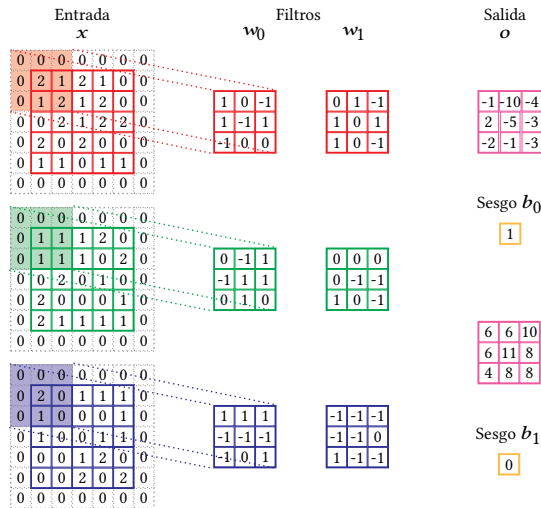


Convolución: entrada 5×5 , filtro 3×3 , paso 2×2 , relleno 1×1 , salida 3×3 .

Capa de convolución (I)



Capa de convolución (II)

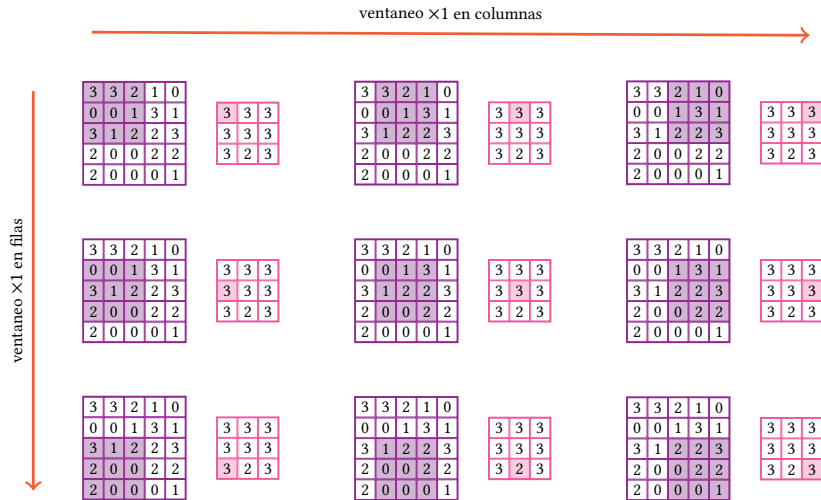


Convolución: entrada $5 \times 5 \times 3$, filtro $3 \times 3 \times 3$, paso 2×2 , relleno 1×1 , salida $3 \times 3 \times 2$.

- Entrada: $C_I \times H_I \times W_I$
- Hiperparámetros:
 - K : número de filtros, profundidad de la salida.
 - F : tamaño del filtro, extensión espacial del filtro.
 - S : paso, cantidad de desplazamiento del filtro.
 - P : relleno, cantidad de aumento de ceros.
- Salida: $C_O \times H_O \times W_O$
 - $H_O = \frac{(H_I - F + 2P)}{S} + 1$
 - $W_O = \frac{(W_I - F + 2P)}{S} + 1$
 - $C_O = K$

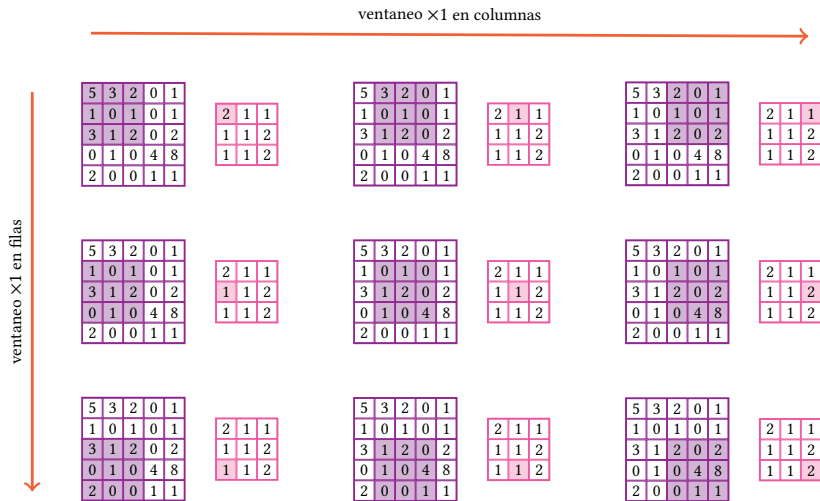
- ¿Cuales serían las dimensiones del bloque de salida para una capa convolucional con siguientes características?
 - Entrada: $1 \times 28 \times 28$
 - Número de filtros: 4
 - Tamaño del filtro: 3
 - Salto: 1
 - Relleno: 1
 - $H_O = \frac{(H_I - F + 2P)}{S} + 1$
 - $W_O = \frac{(W_I - F + 2P)}{S} + 1$
 - $O = K$

Capa de submuestreo máximo



Muestreo máximo: entrada 5×5 , paso 1×1 , salida 3×3 .

Capa de submuestreo promedio



Muestreo máximo: entrada 5×5 , paso 1×1 , salida 3×3 .

- Entrada: $C_I \times H_I \times W_I$
- Hiperparámetros:
 - K : tamaño del filtro, extensión espacial.
 - S : paso, cantidad de desplazamiento del filtro.
- Salida: $C_O \times H_2 \times W_2$
 - $H_O = \frac{H_I - F}{S} + 1$
 - $W_O = \frac{W_I - F}{S} + 1$
 - $C_O = C_I$

- ¿Cuales serían las dimensiones del bloque de salida para una capa de muestreo con siguientes características?

- Entrada: $4 \times 28 \times 28$

- Tamaño del filtro: 2×2

- Salto: 2×2

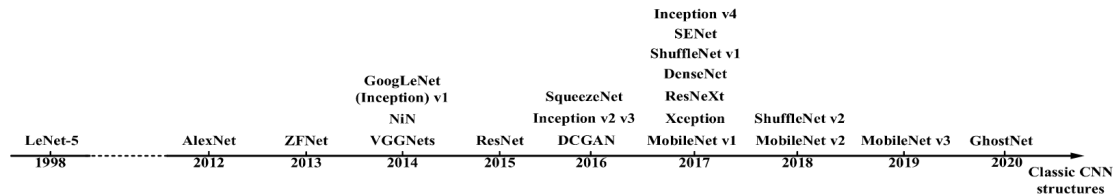
- $H_O = \frac{H_I - F}{S} + 1$

- $W_O = \frac{W_I - F}{S} + 1$

- $C_O = C_I$

Arquitecturas

Cronología de las arquitecturas convolucionales



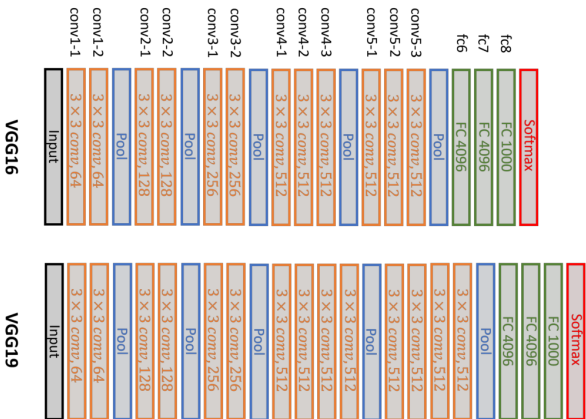


Imagen tomada de diapositivas de <http://datahacker.rs/deep-learning-vgg-16-vs-vgg-19/>

Factorización de convoluciones

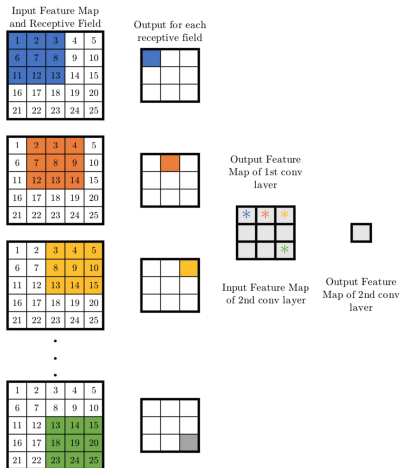


Imagen tomada de <https://inblog.in/Data-Science-Interview-Questions-30-days-of-Interview-Preparation-Day-14-asi4Xq3QNc>



GoogLeNet (Inception v1)

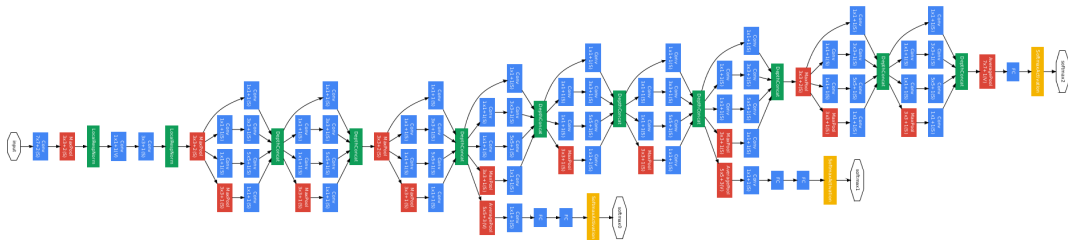


Imagem tomada de <https://inblog.in/Data-Science-Interview-Questions-30-days-of-Interview-Preparation-Day-14-asi4Xq3QNc>

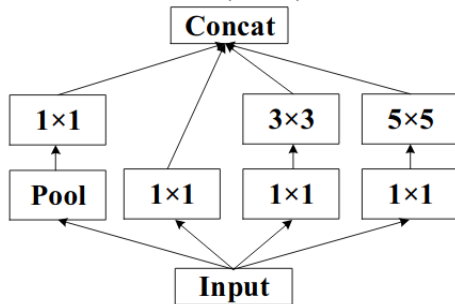


Imagem tomada de <https://research.googleblog.com/2016/03/train-your-own-image-classifier-with.html>

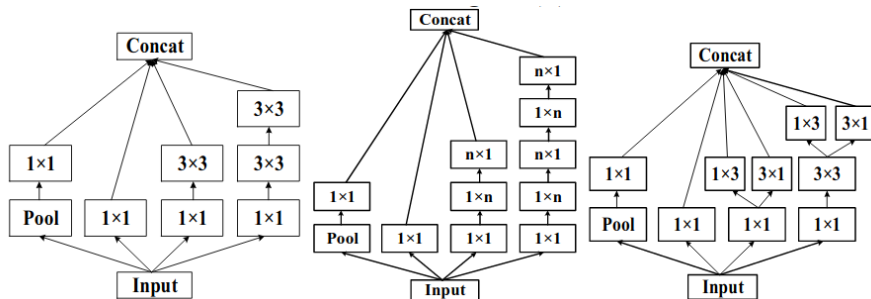


Imagen tomada de <https://research.googleblog.com/2016/03/train-your-own-image-classifier-with.html>

- Inception v3: Utiliza factorización de convoluciones de 5×5 y 3×3 . La entrada de la red es de 299×299 .
- Inception v4: Utiliza más módulos Inception en comparación con la versión anterior.

Szegedy et al. *Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning*. 2016.

Szegedy et al. *Rethinking the Inception Architecture for Computer Vision*. 2015.