Aprendizaje profundo

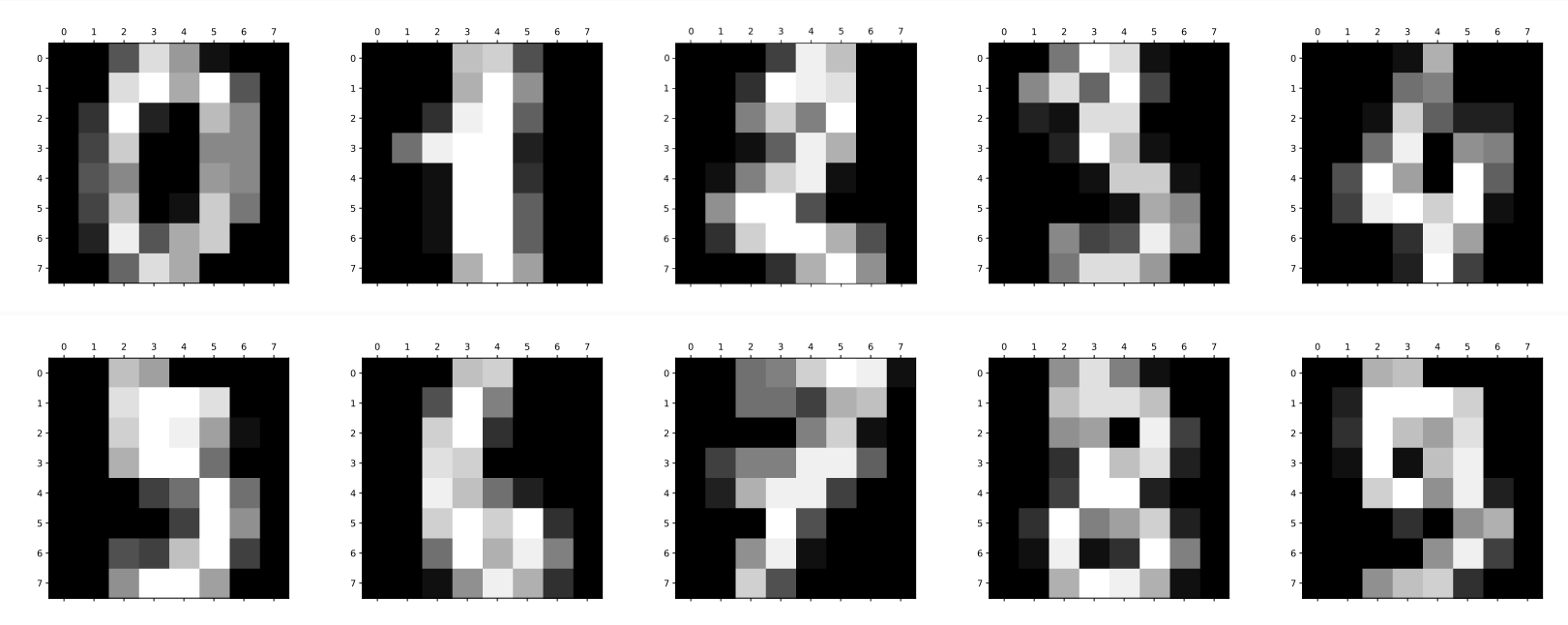
Redes convolucionales

Gibran Fuentes-Pineda

Agosto 2019

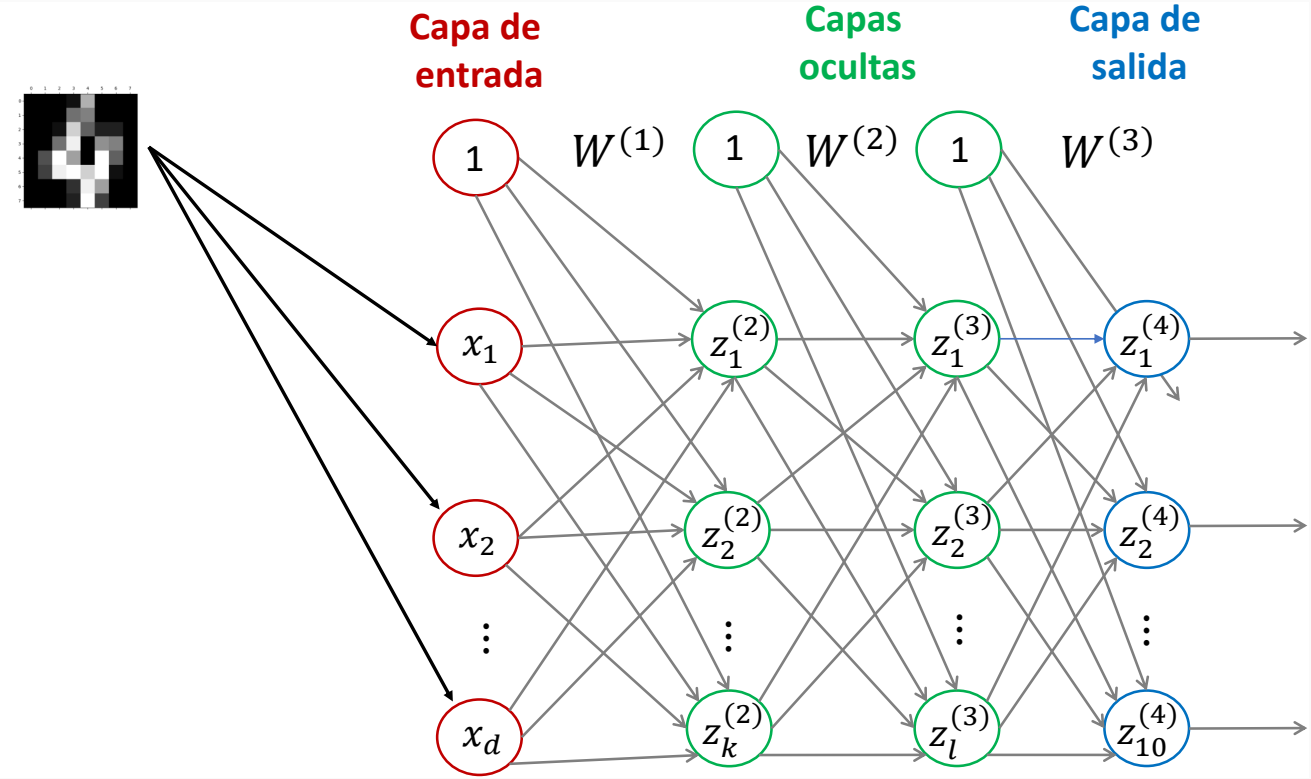
Clasificación de imágenes

- Por ejemplo, clasificar dígitos escritos a mano



Clasificación de imágenes con PMC

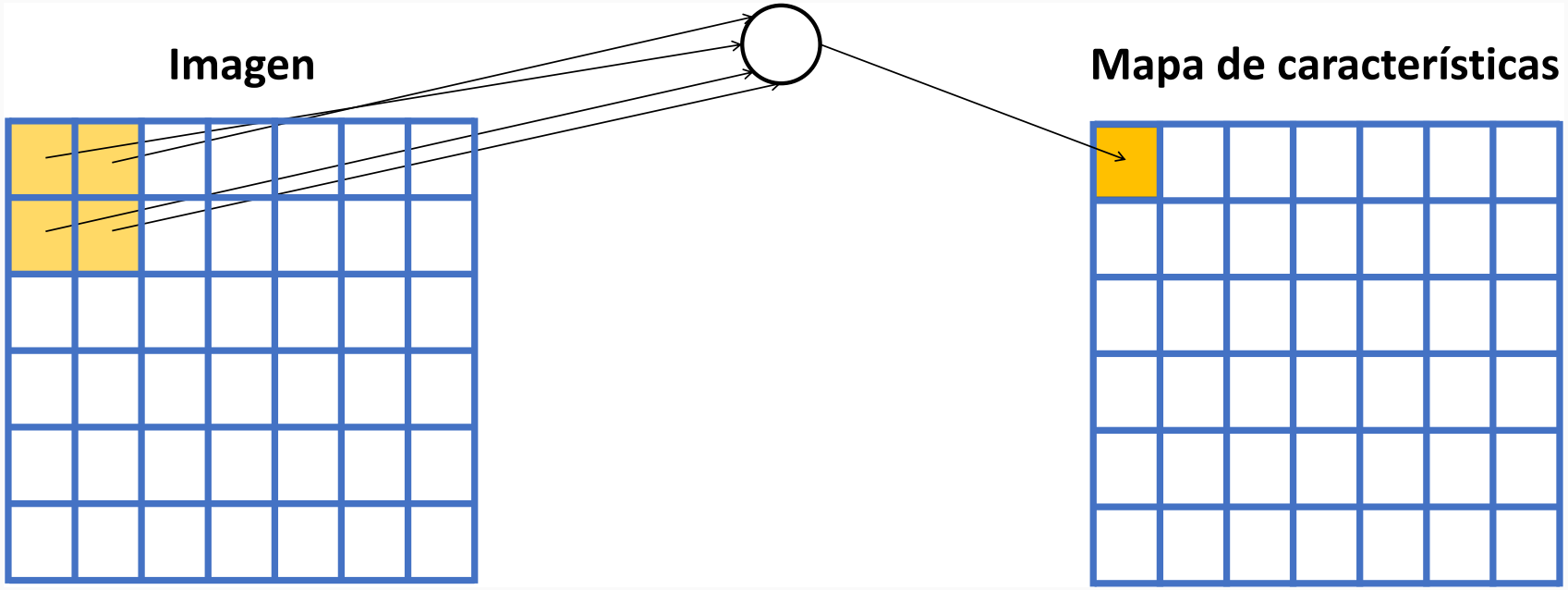
- Redes perceptrón multicapa tienen muchos parámetros
  - Por ej., si tenemos imágenes de 32×32 y 1 red con 1 sóla capa oculta con 10 neuronas tendríamos $(32 \times 32 \times 10) + 1$ pesos



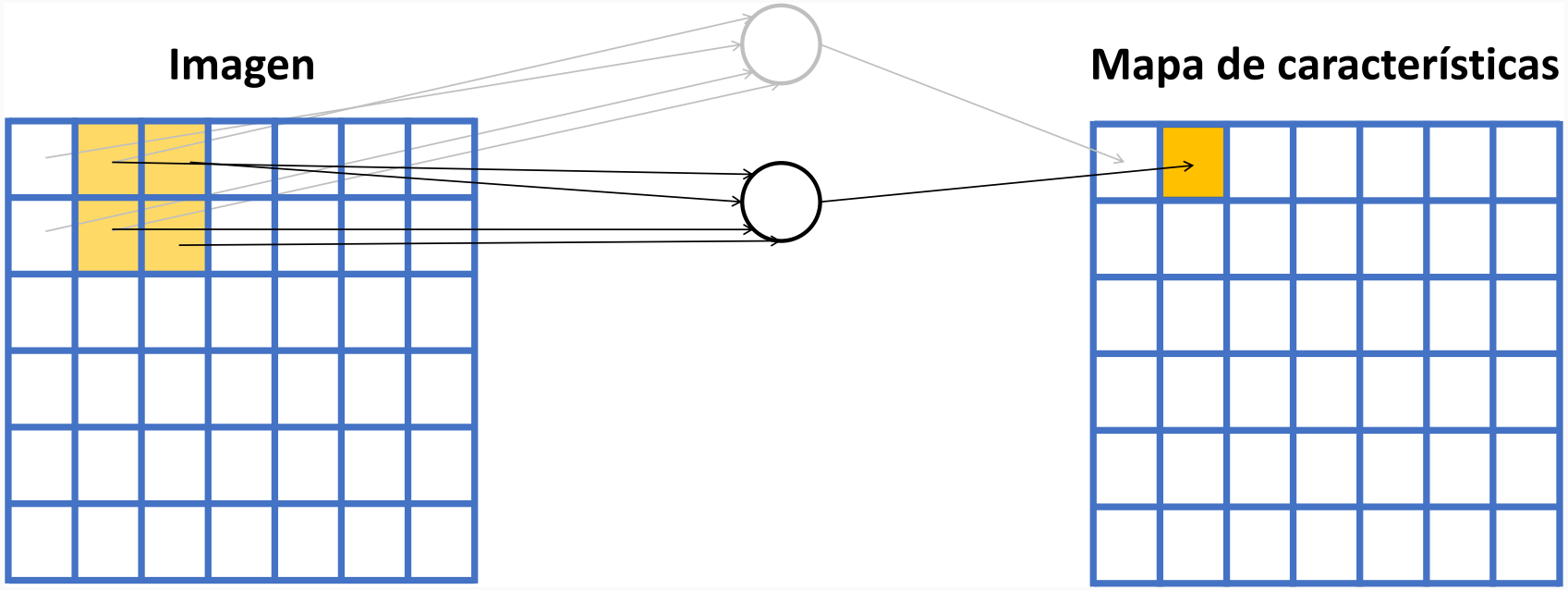
Capa convolucional

- Pueden verse como un caso especial de una capa densa con 2 variaciones
  1. Conectividad local
  2. Pesos compartidos

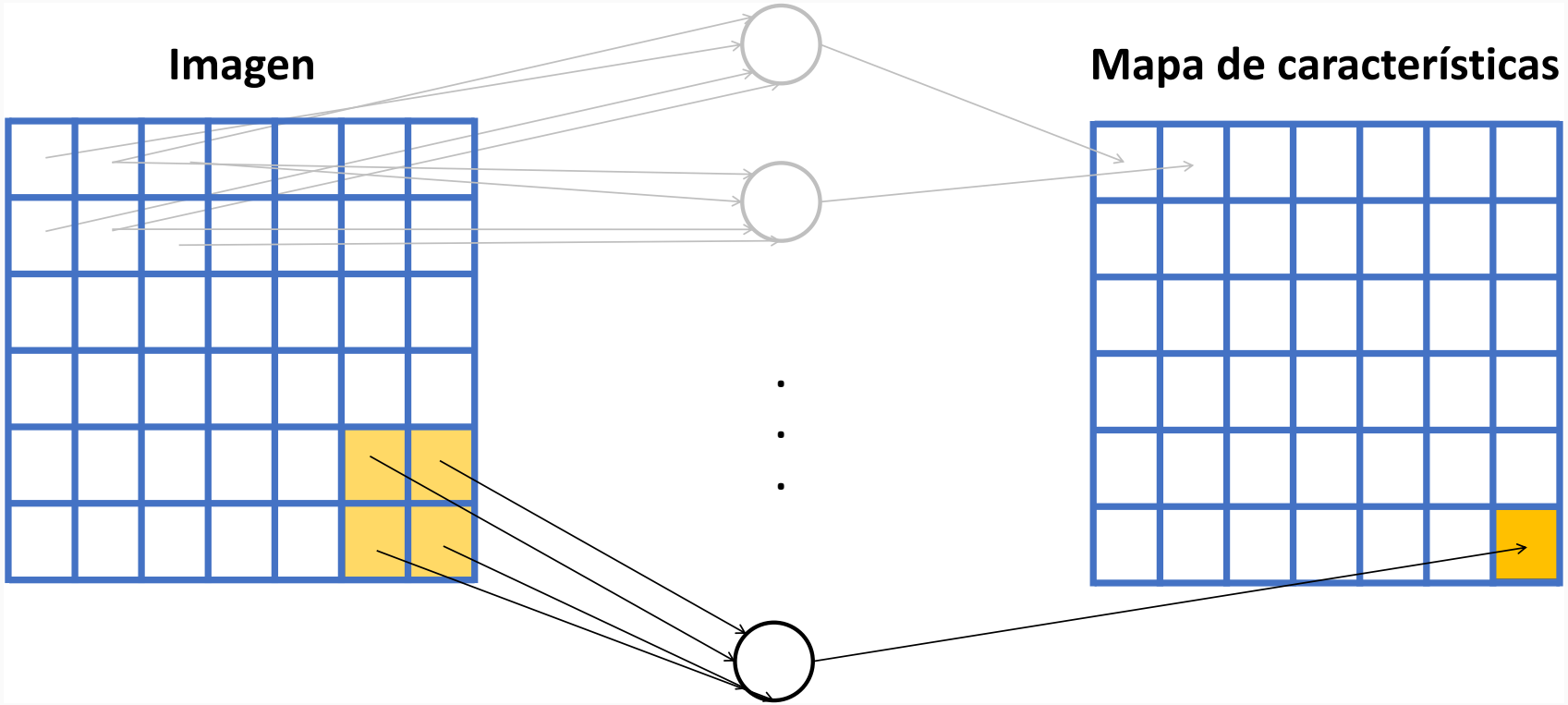
Conectividad local (1)



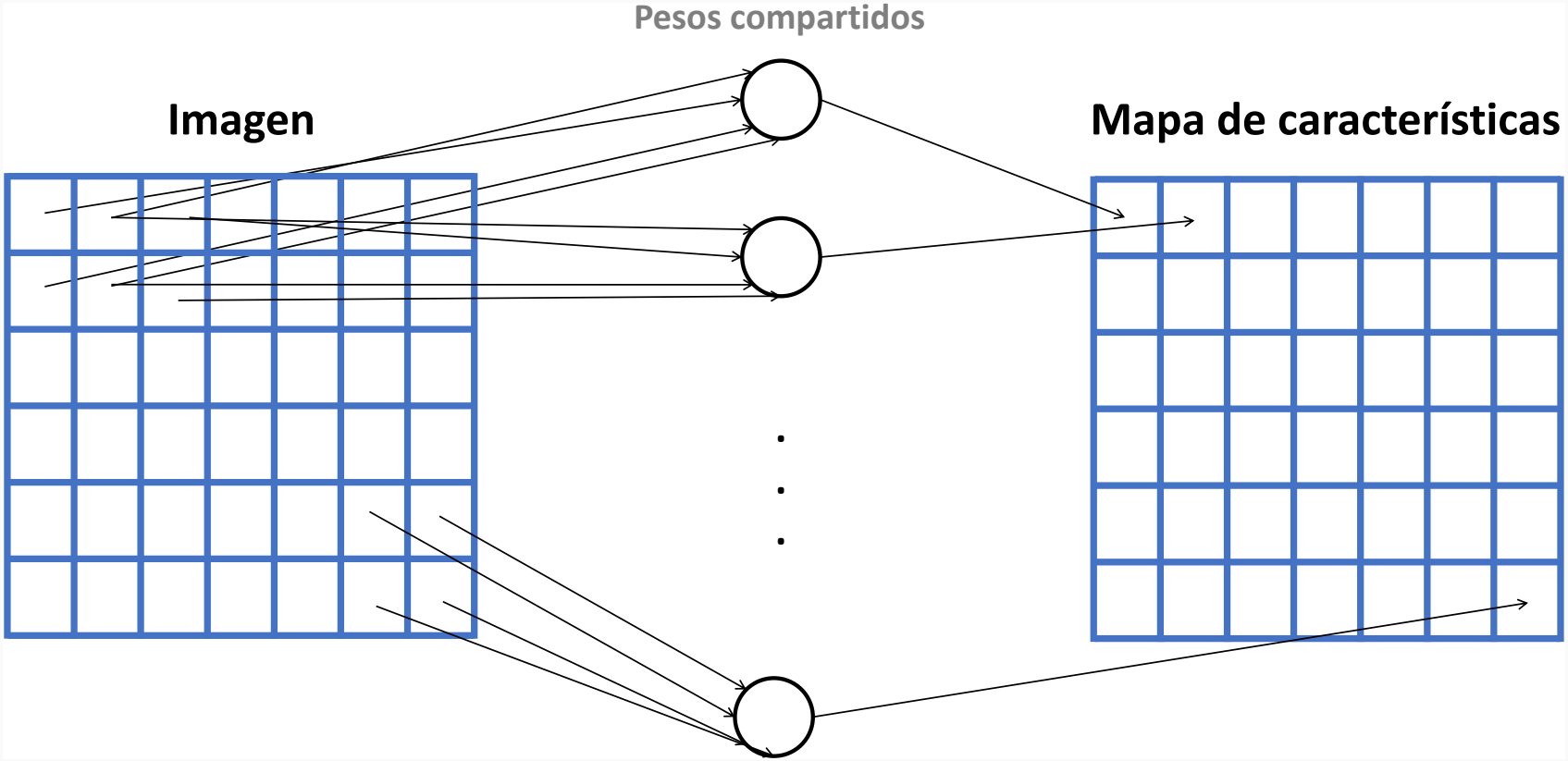
Conectividad local (2)



Conectividad local (3)



Pesos compartidos



Operación de convolución

- Convolución en 1 dimensión
  - Continua $s(t) = (x * k)(t) = \int_m x(m)k(t-m)dm$
  - Discreta $s[i] = (x * k)[i] = \sum_{m=-\infty}^{\infty} x[m]k[i-m]$

- Convolución en 2 dimensiones (discreta)
  - $S[i,j] = (I * K)[i,j] = \sum_m \sum_n I[m,n]K[i-m, j-n]$
  - $S[i,j] = (K * I)[i,j] = \sum_m \sum_n I[i-m, j-n]K[m,n]$

- Correlación cruzada (discreta)
  - $S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(i+m, j+n)K(m,n)$

Operación de convolución

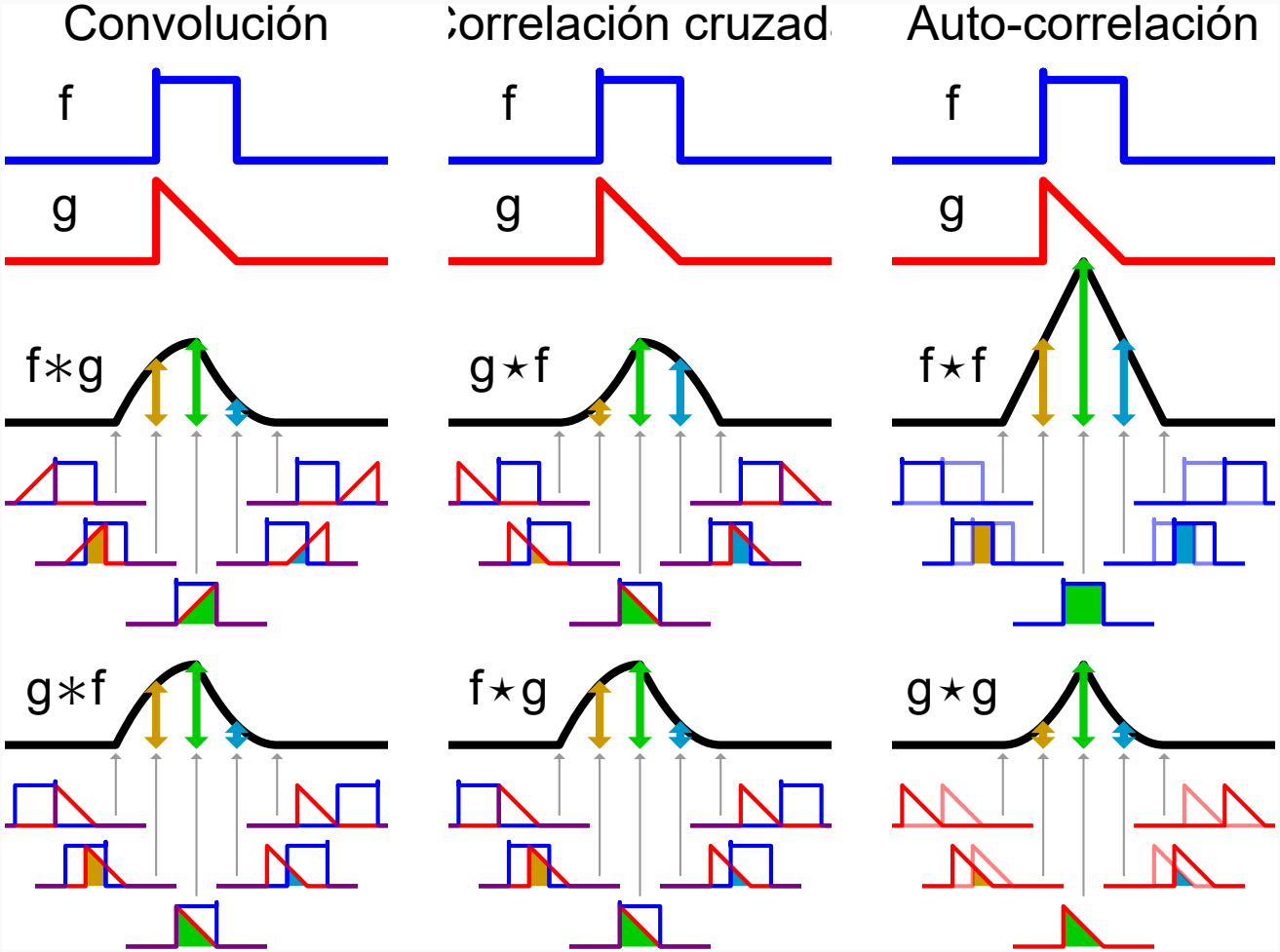
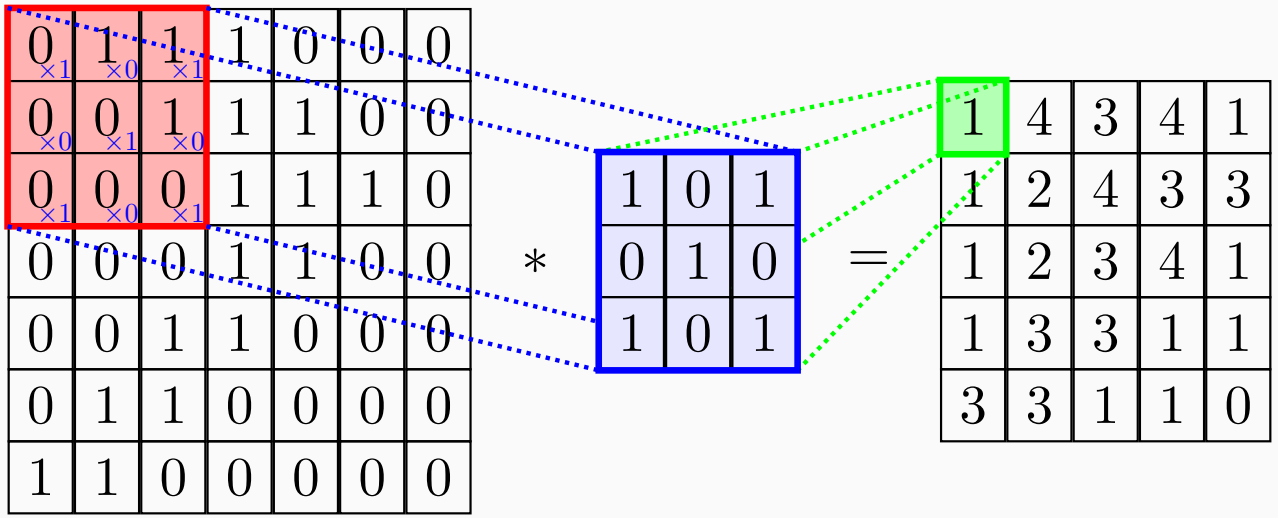
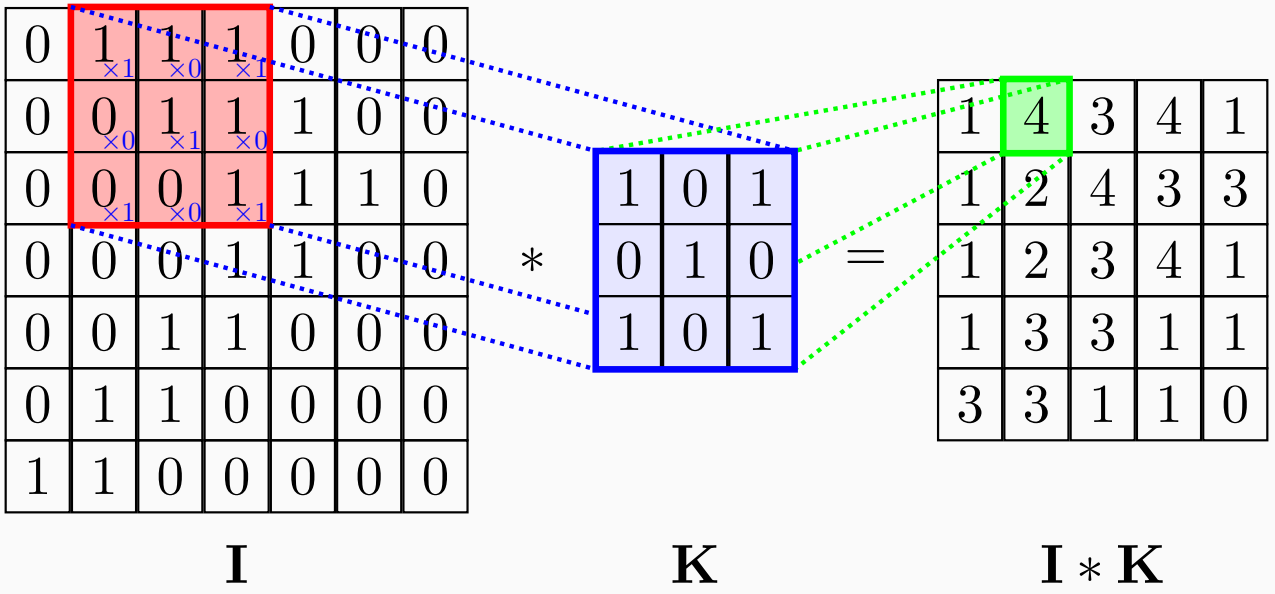


Imagen tomada de Wikipedia (Convolution)

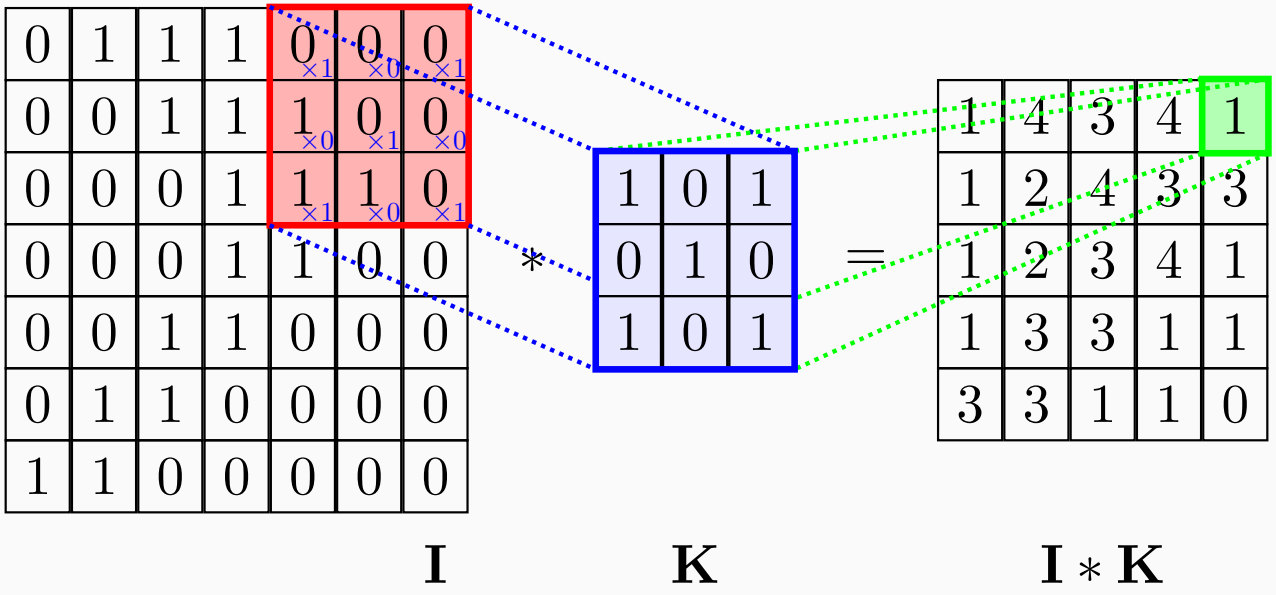
Convolución de una imagen con un filtro (1)



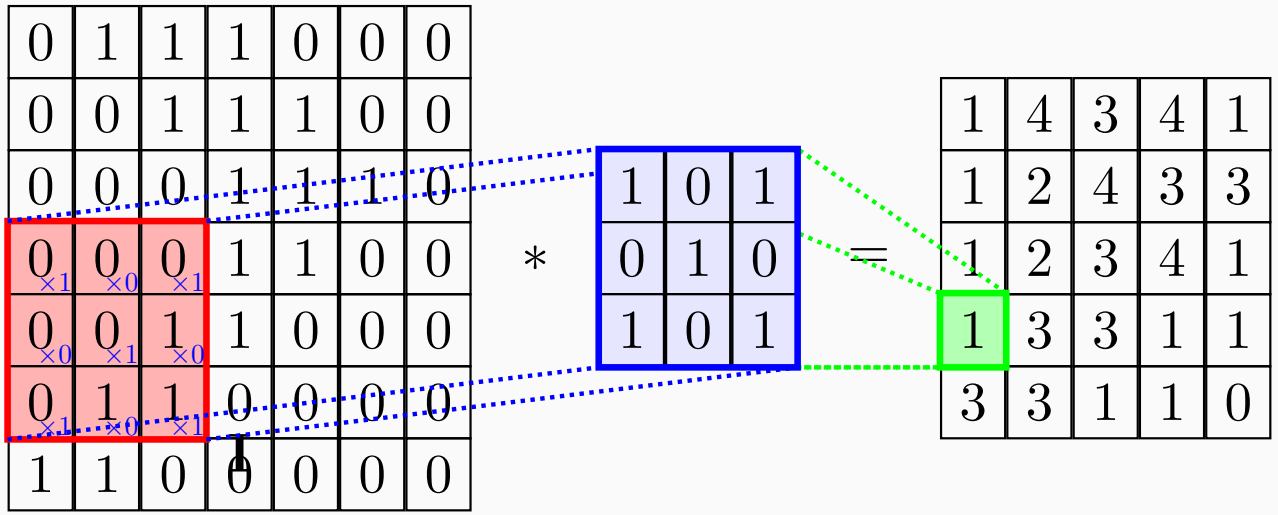
Convolución de una imagen con un filtro (2)



Convolución de una imagen con un filtro (3)



Convolución de una imagen con un filtro (4)



Mapas de activación

- Mapas de activación pueden verse como volúmenes

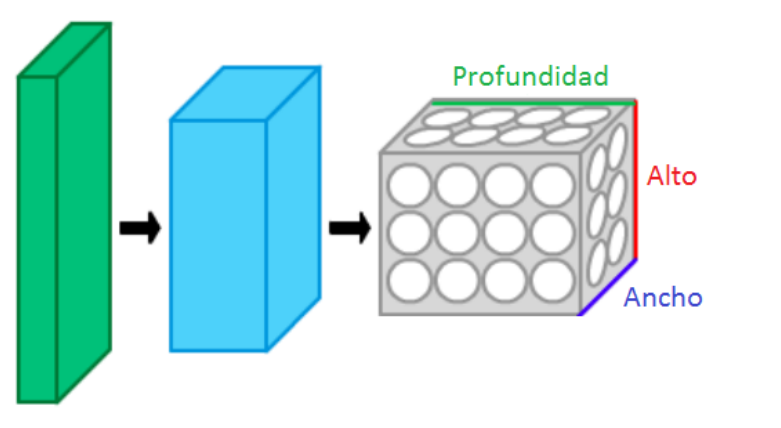


Imagen tomada de Wikipedia (Convolutional Neural Network)

- Dimensiones están dadas por $H^{\{\ell\}} \times W^{\{\ell\}} \times N_{filters}$, donde

$$H^{\{\ell\}} = \frac{H^{\{\ell-1\}} + 2P^{\{\ell-1\}} - F_H^{\{\ell\}}}{S^{\{\ell\}}} + 1$$

$$W^{\{\ell\}} = \frac{W^{\{\ell-1\}} + 2P^{\{\ell-1\}} - F_W^{\{\ell\}}}{S^{\{\ell\}}} + 1$$

Submuestreo o decimación

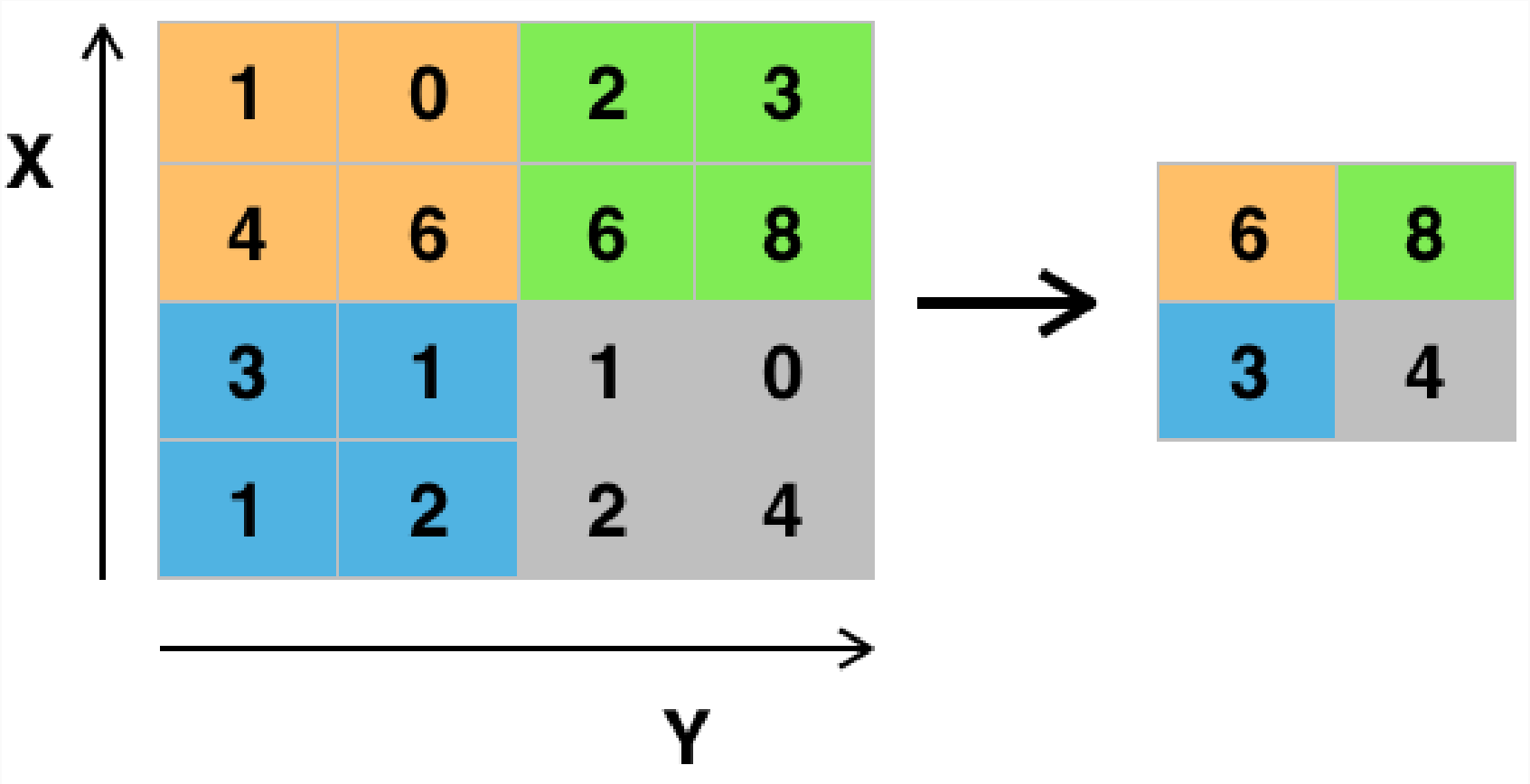


Imagen tomada de Wikipedia (Convolutional Neural Network)

Retro-propagación en redes convolucionales

- Hacia adelante
  - Se aplican las operaciones de convolución con correspondiente activación y decimación
- Hacia atrás
  - En la convolución cada neurona actualiza los gradientes por separado y al final se suman para actualizar los pesos compartidos
  - En la decimación se actualiza sólo la neurona ganadora (max-pooling)

Convolución de 1×1

- Funciona como un tipo de decimación en profundidad

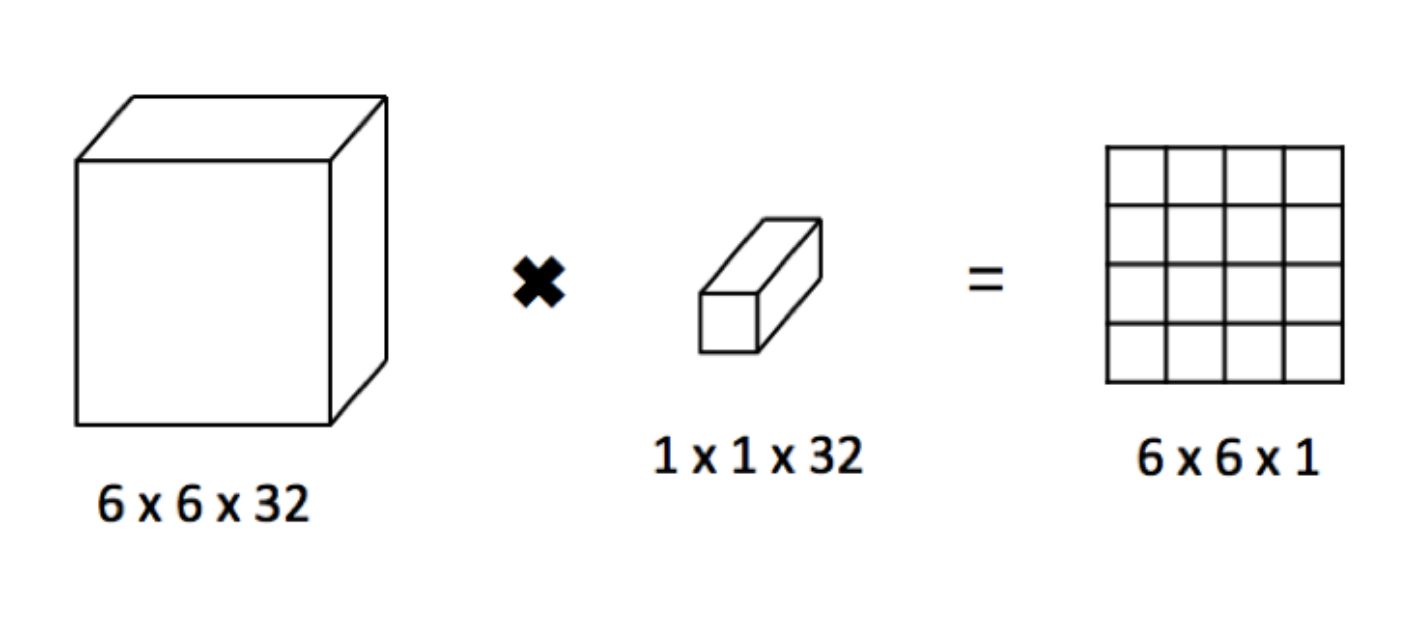


Imagen tomada de

https://medium.com/machine-learning-bites/deeplearning-series-convolutional-neural-networks-a9c2f2ee1524

Arquitecturas de redes convolucionales: LeNet

- Red poco profunda inspirada en la propuesta originalmente por LeCun et al. 1998

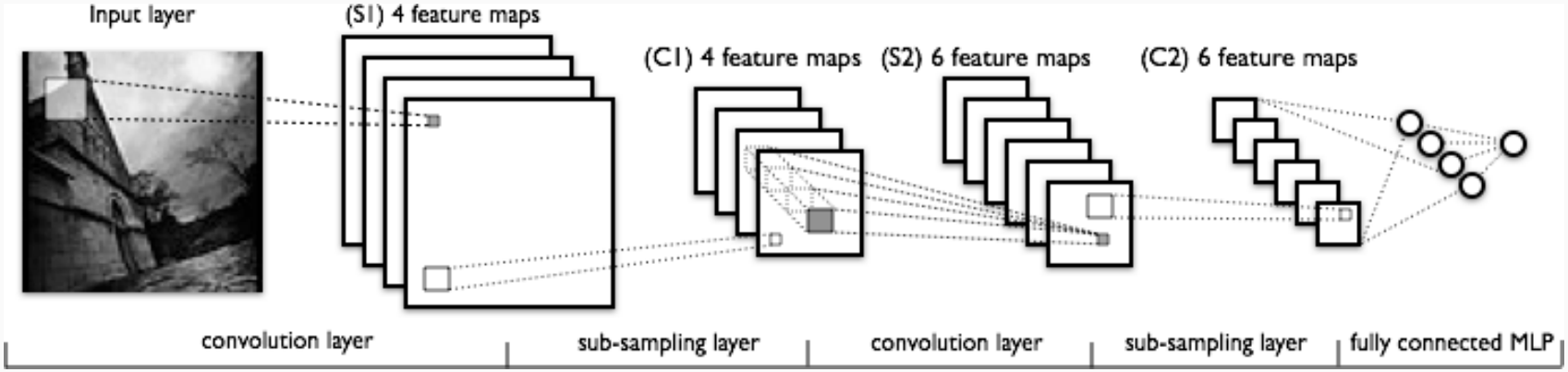


Imagen tomada de `http://deeplearning.net/tutorial/lenet.html`

Arquitecturas de redes convolucionales: AlexNet

- Usa función de activación ReLU
- Entrenamiento con version optimizada para 2 GPUs
- Normaliza respuestas
- Submuestreo con traslape

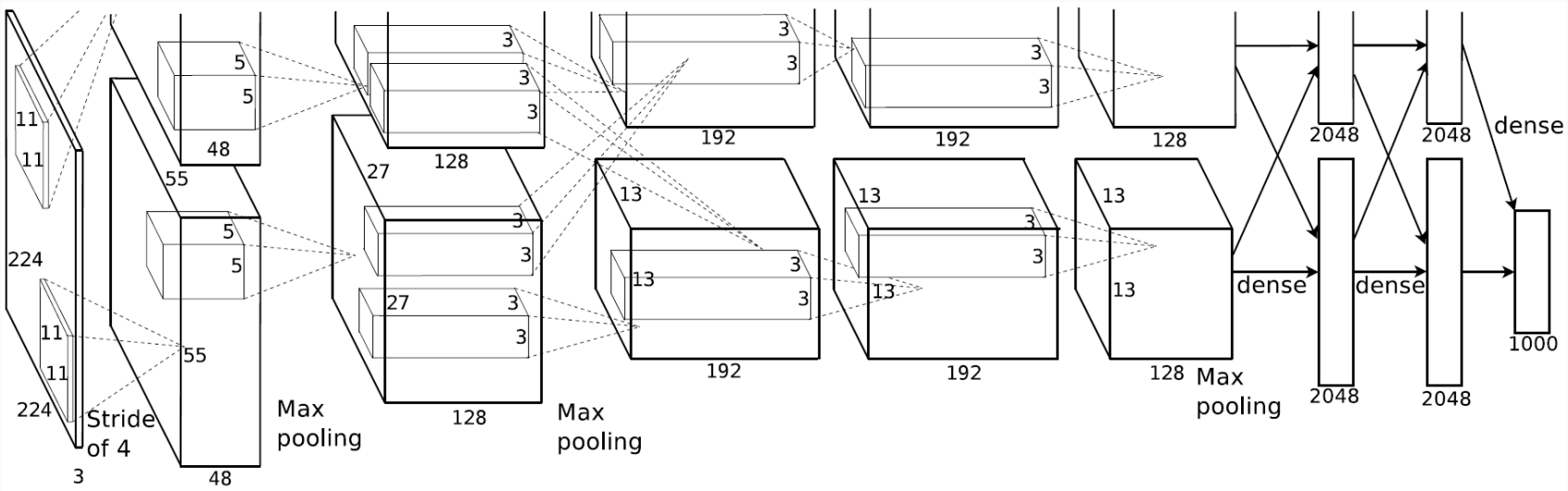


Imagen tomada de Krizhevsky et al. *ImageNet Classification with Deep Convolutional Neural Networks*, 2012

Arquitecturas de redes convolucionales: ZFNet

- Muy similar a AlexNet pero usa filtros más pequeños con menor desplazamiento

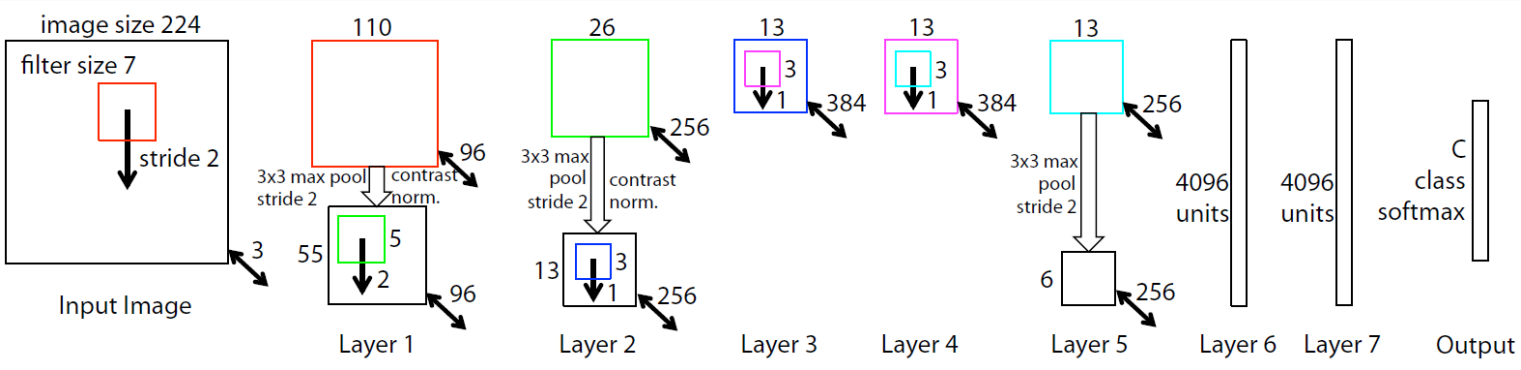


Imagen tomada de Zeiler and Fergus. *Visualizing and Understanding Convolutional Networks*, 2014

Arquitecturas de redes convolucionales: VGGNet

- 19 capas
- Filtros de 3×3 con desplazamientos de 1
- Max-pooling de 2×2 con desplazamientos de 2

Imagen tomada de diapositivas de Simonyan (ILSVRC Workshop 2014)

Arquitecturas de redes convolucionales: GoogleNet

- 22 capas
- Utiliza bloque de capas convolucionales paralelas cuyas salidas se concatenan

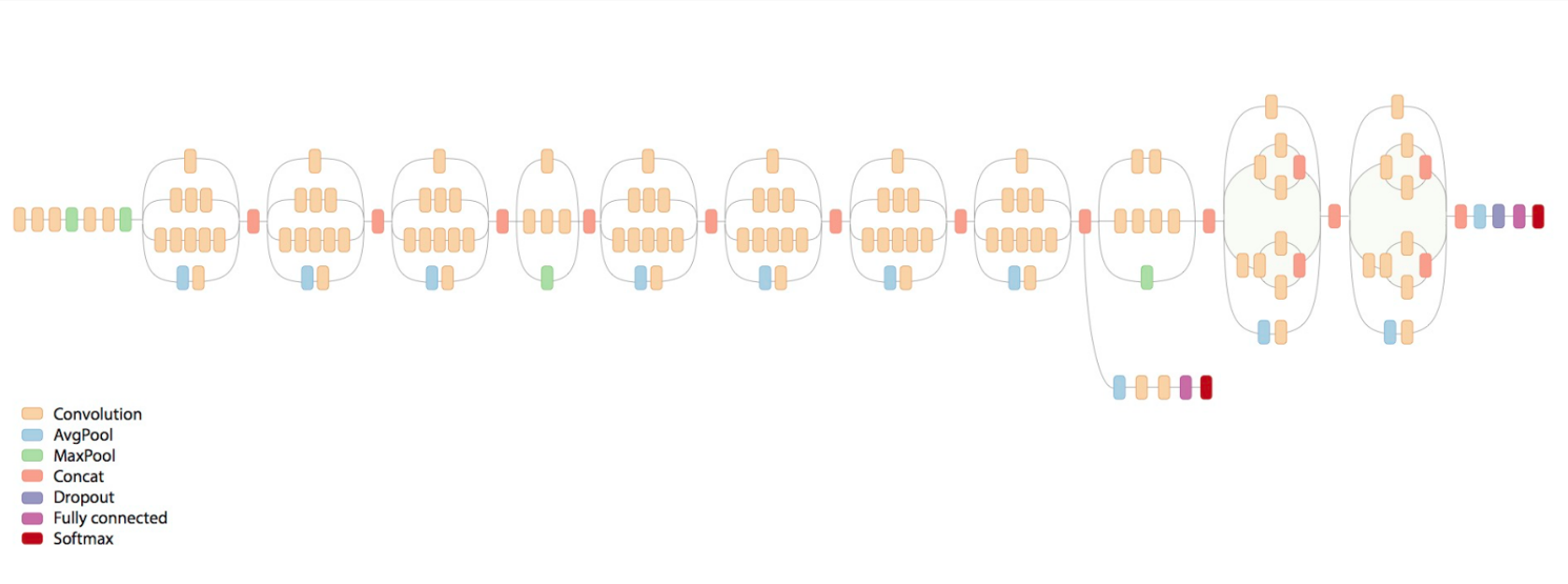
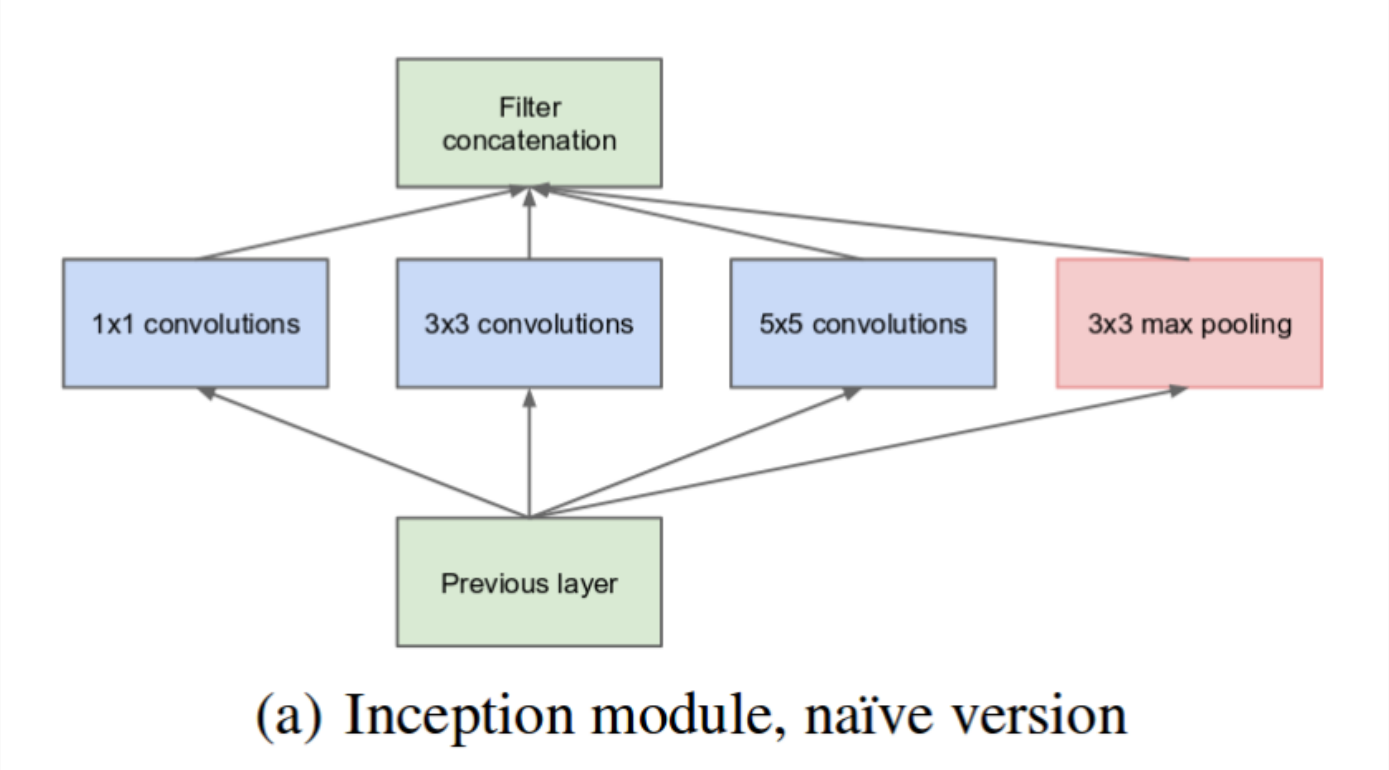


Imagen tomada de `https://research.googleblog.com/2016/03/train-your-own-image-classifier-with.html`

Bloque Inception



(a) Inception module, naïve version

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

Sobre profundidad en redes convolucionales (1)

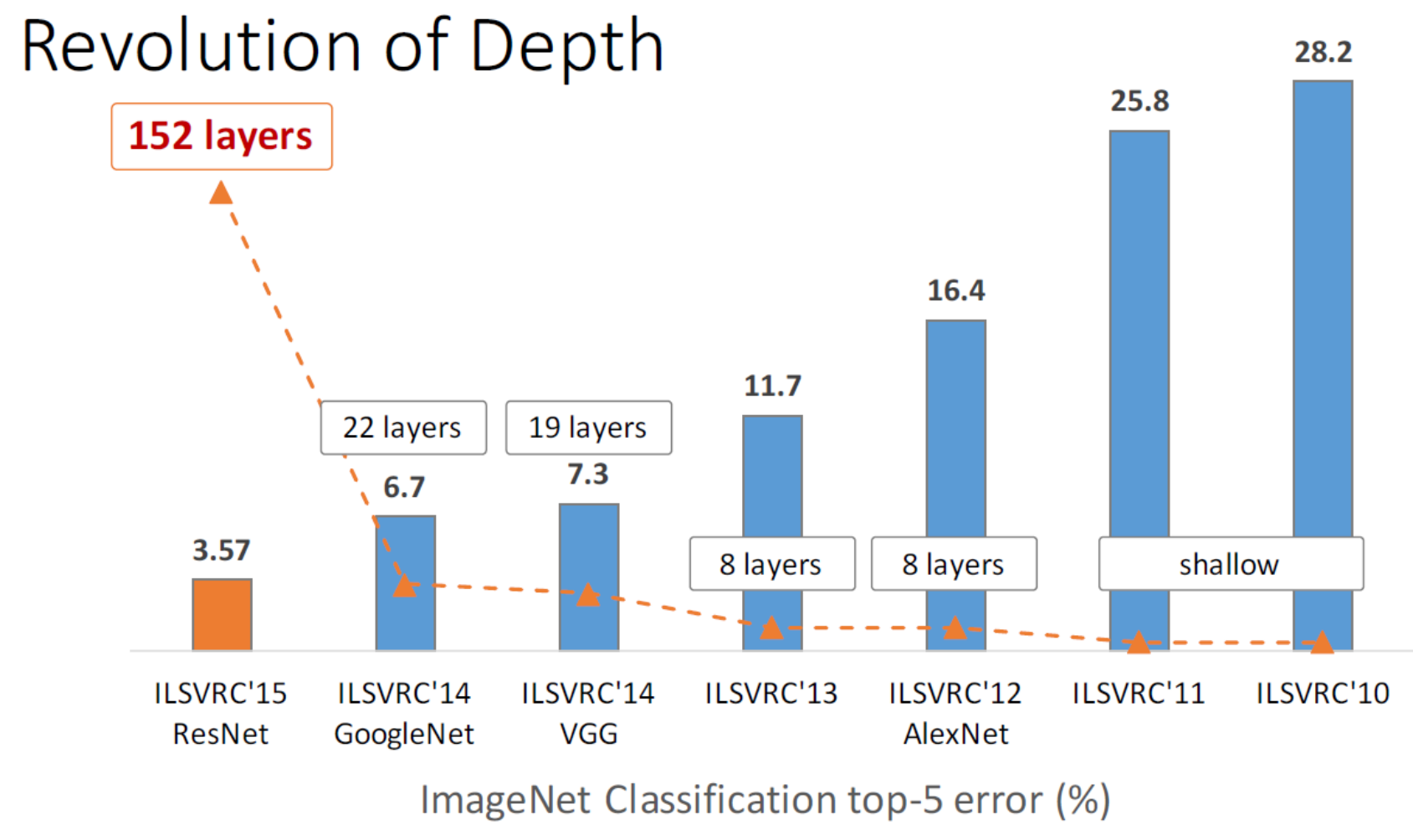


Imagen tomada de diapositivas de Kaiming He (ICML 2016)

Sobre profundidad en redes convolucionales (2)

- Problemas con el desvanecimiento y explosión de respuestas (hacia adelante) y gradientes (hacia atrás)

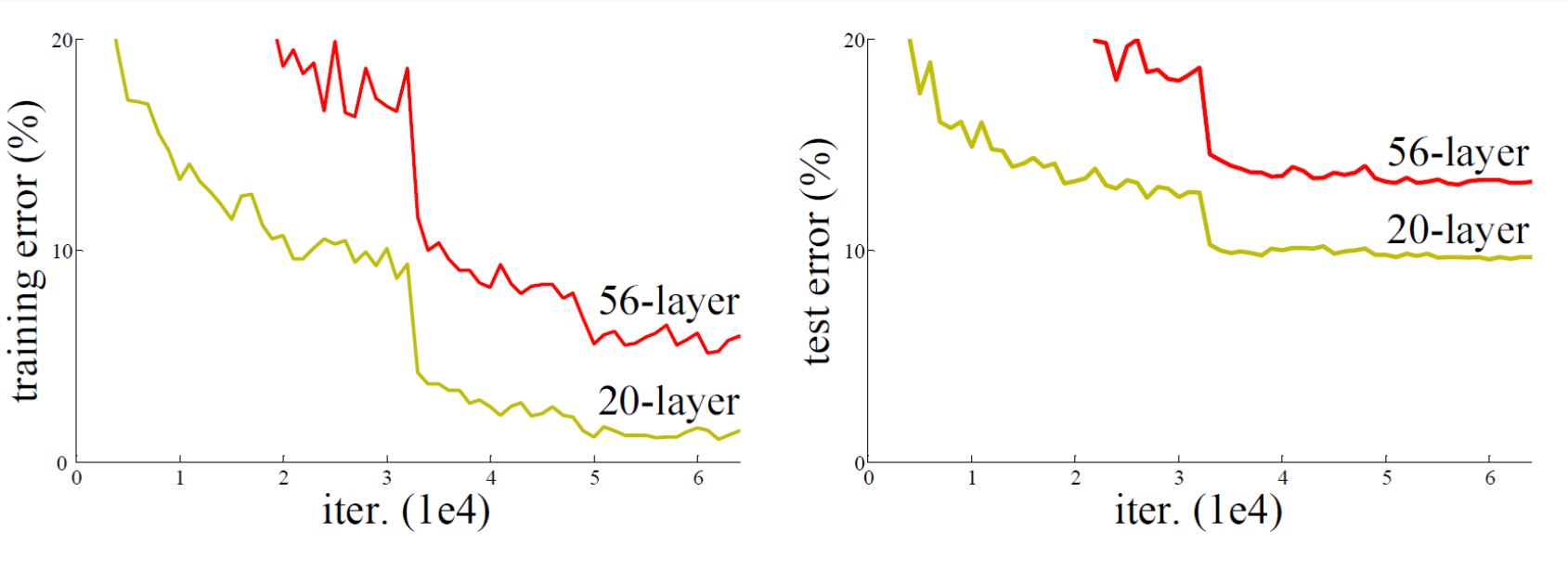
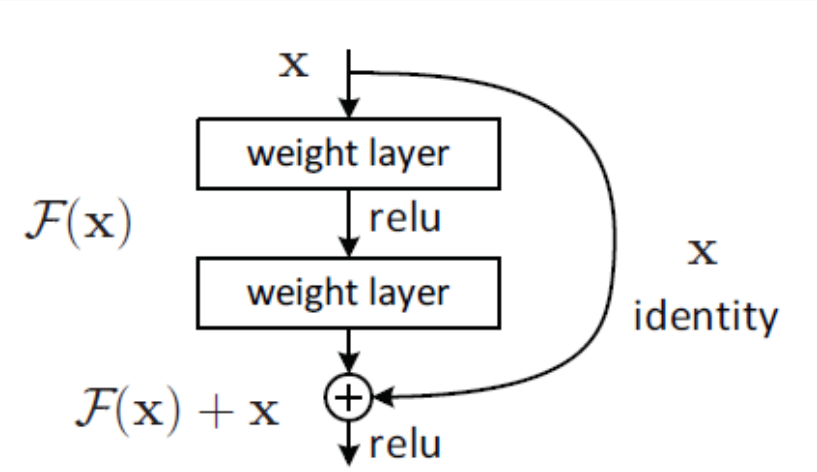


Imagen tomada de He et al. *Deep Residual Learning for Image Recognition* , 2015

Entrenando redes más profundas: ResNet (1)

- Reformula los mapeos que se desean aprender a residuales
- Más fácil de optimizar los residuales que el mapeo original



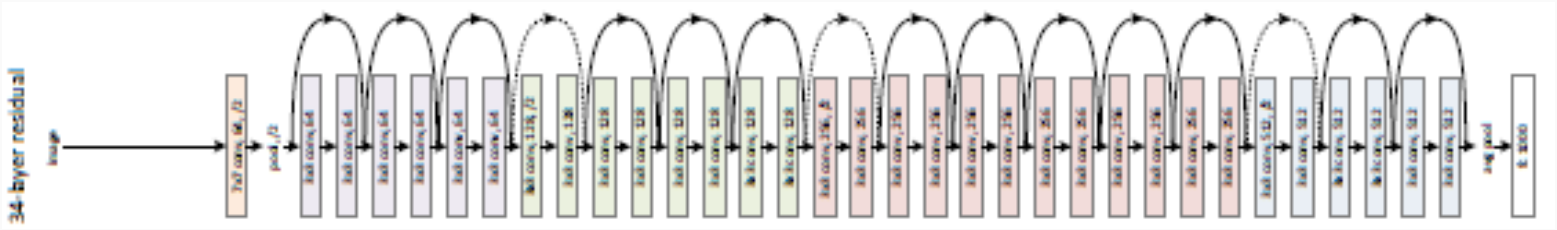


Imagen tomada de He et al. *Deep Residual Learning for Image Recognition* , 2015

Entrenando redes más profundas: ResNet (2)

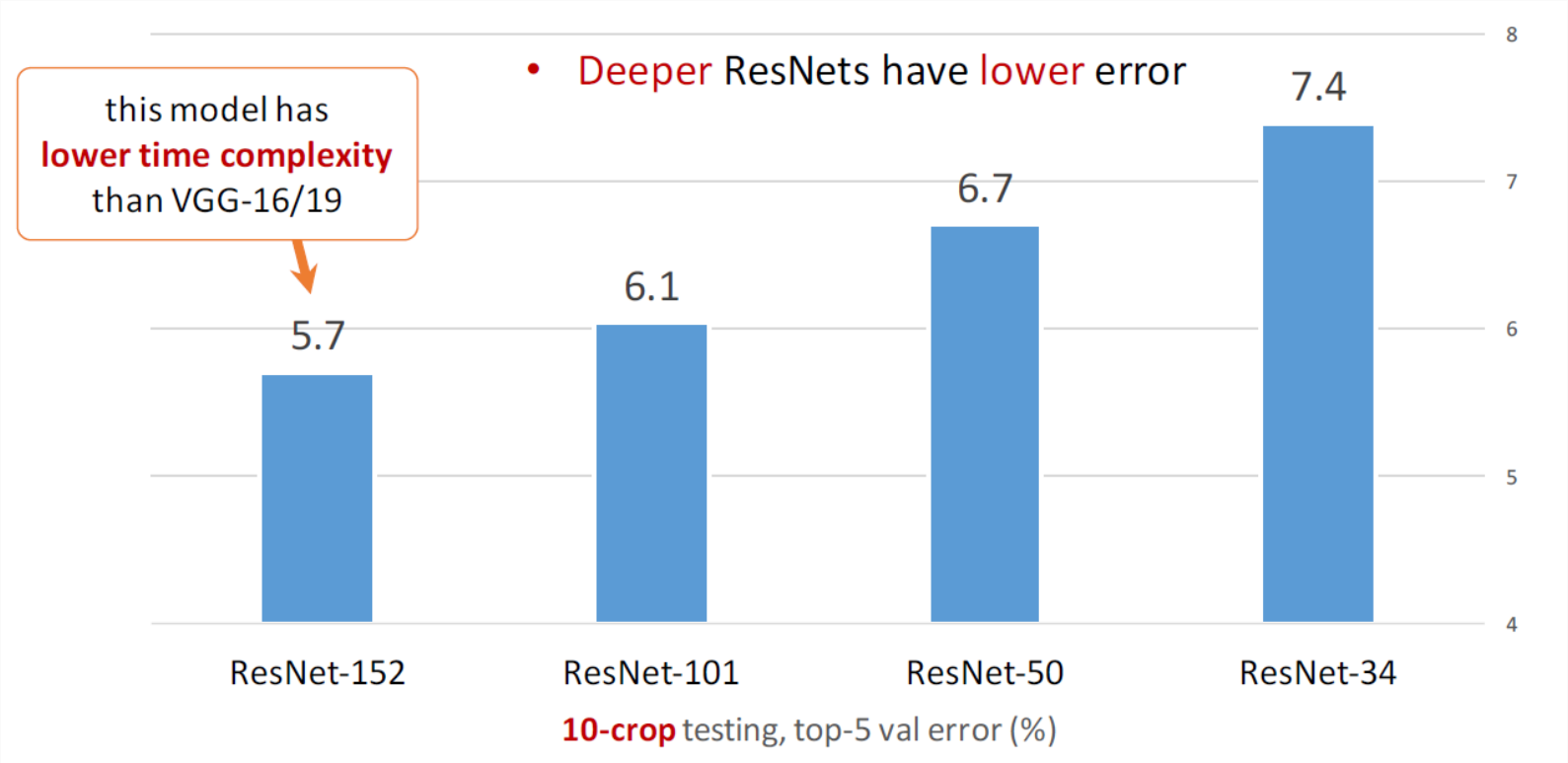


Imagen tomada de diapositivas de Kaiming He (ICML 2016)

Redes convolucionales 1D

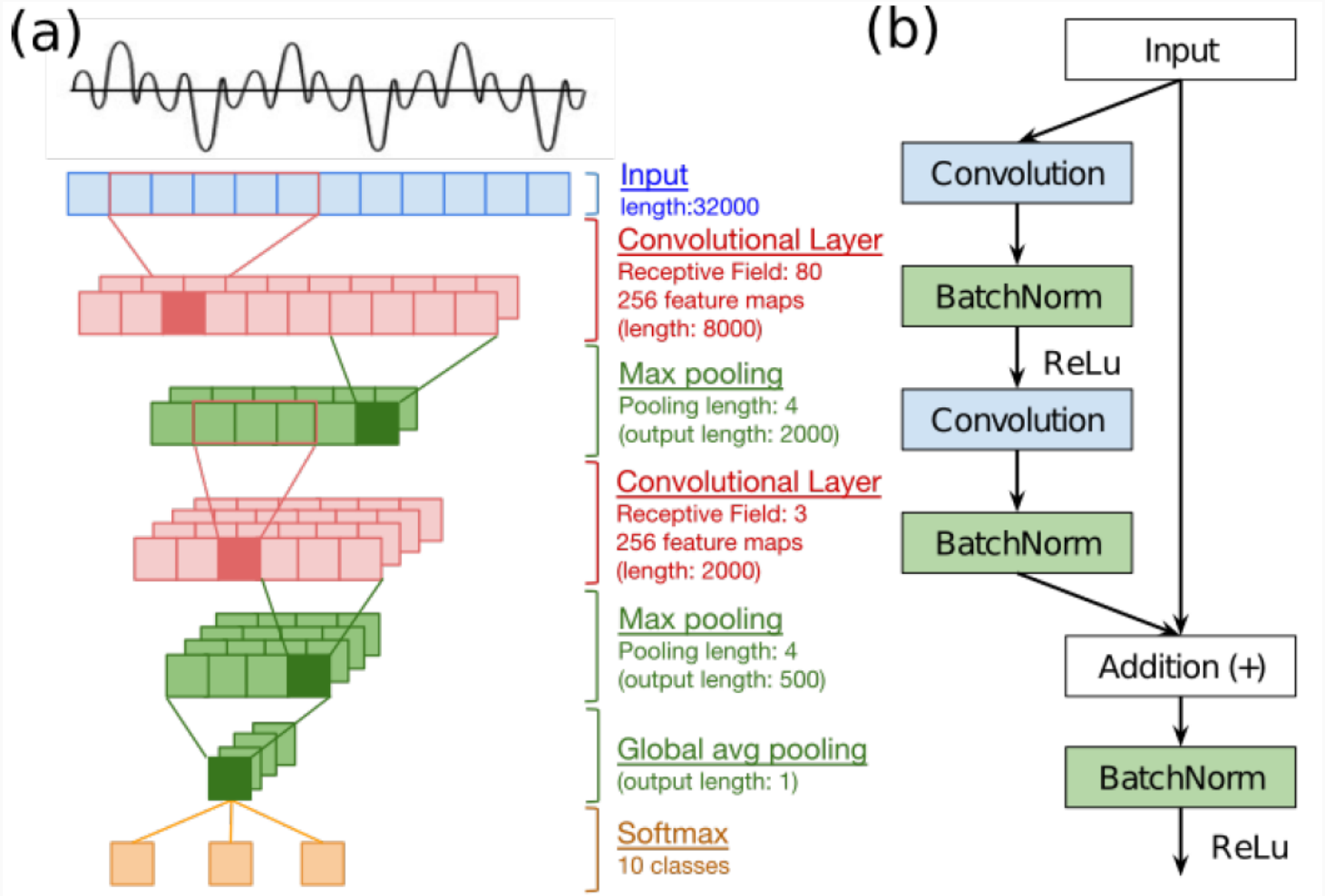


Imagen tomada de Dai et al. *Very Deep Convolutional Neural Networks for Raw Waveforms*, 2016

Redes convolucionales 3D (1)

- Respuesta en redes convolucionales 2D en una imagen

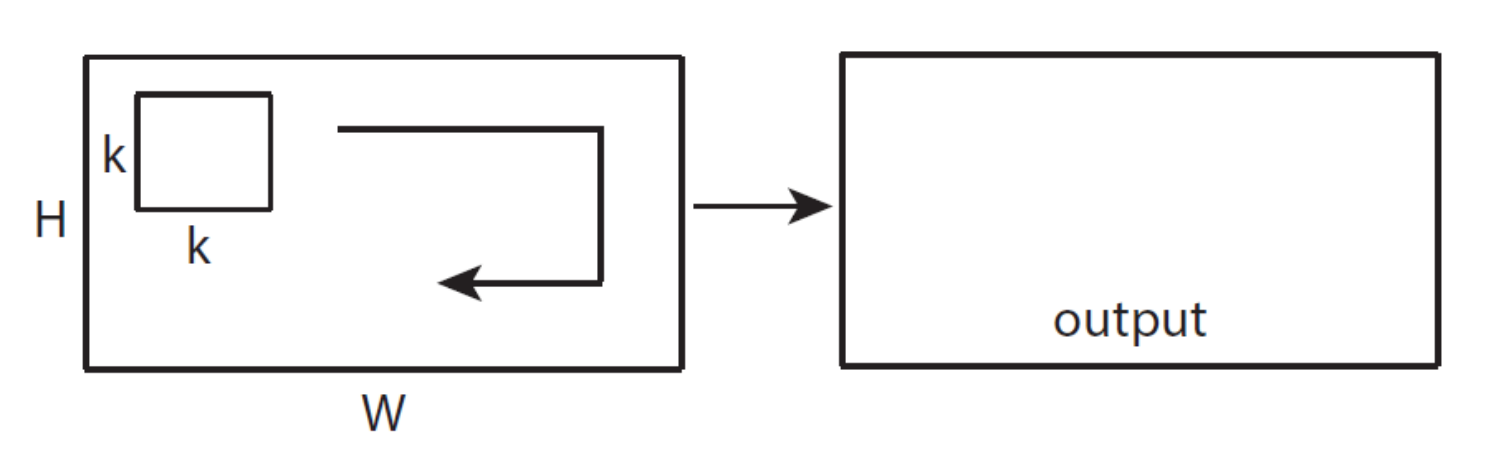


Imagen tomada de Tran et al. *Learning Spatiotemporal Features with 3D Convolutional Networks*, 2015

Redes convolucionales 3D (2)

- Respuesta en redes convolucionales 2D en varias imágenes

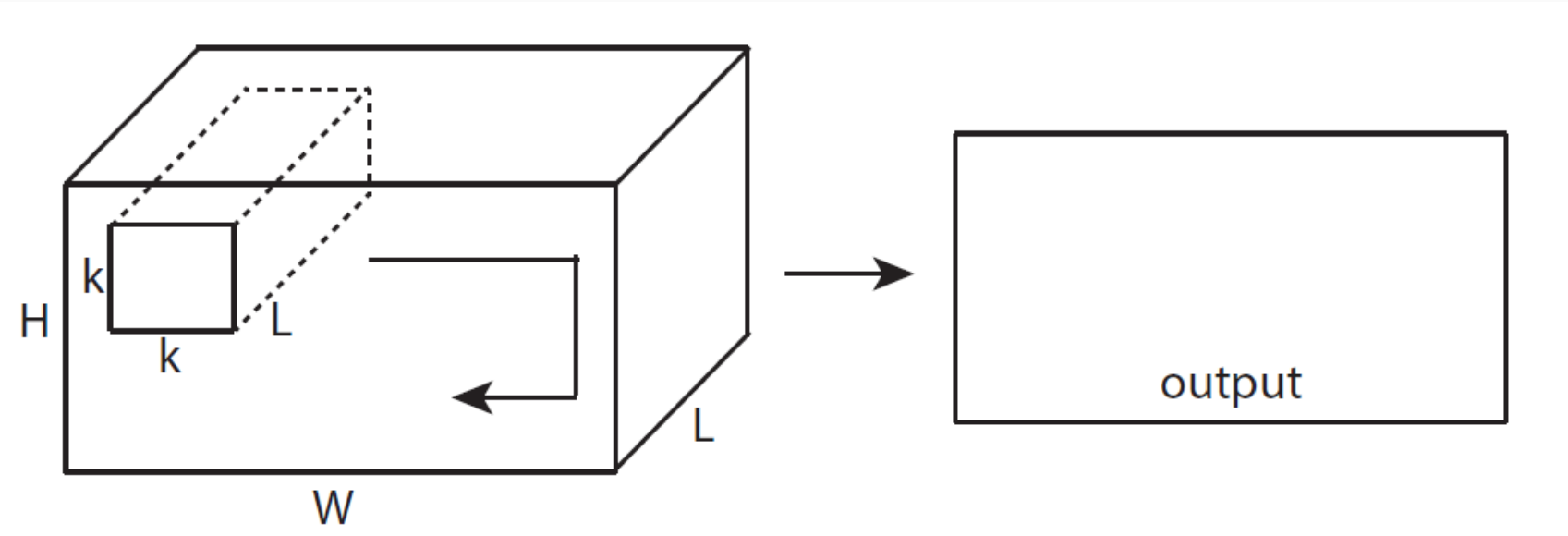


Imagen tomada de Tran et al. *Learning Spatiotemporal Features with 3D Convolutional Networks*, 2015

Redes convolucionales 3D (3)

- Respuesta en redes convolucionales 3D

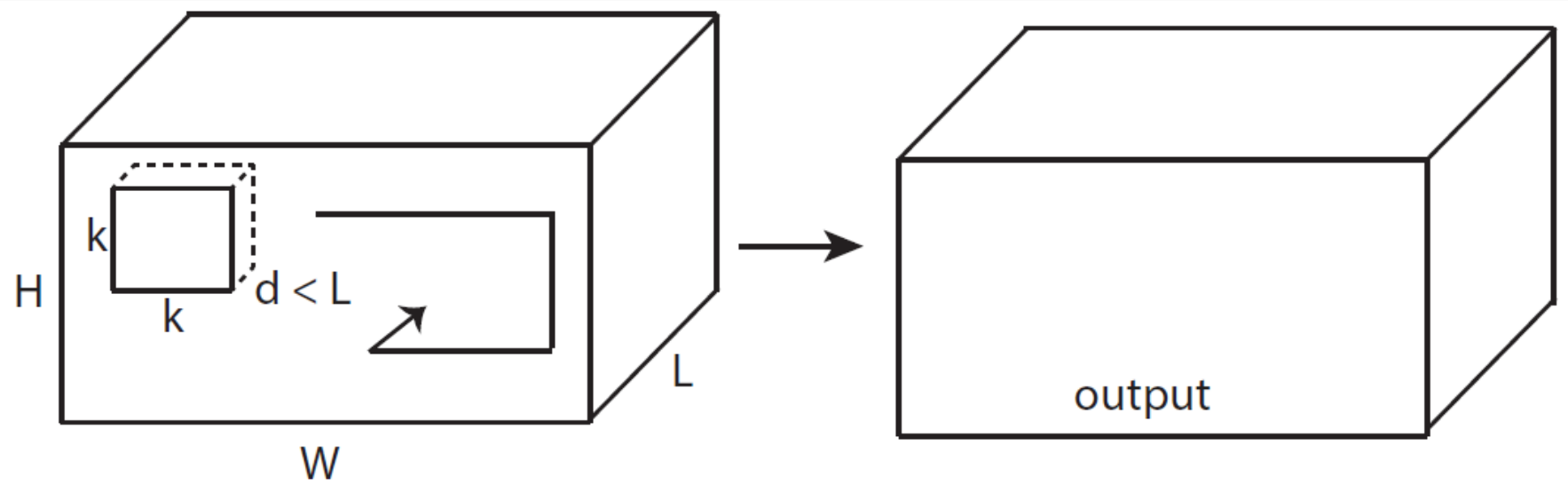


Imagen tomada de Tran et al. *Learning Spatiotemporal Features with 3D Convolutional Networks*, 2015

Filtros separables

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

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

Convolución separable espacialmente

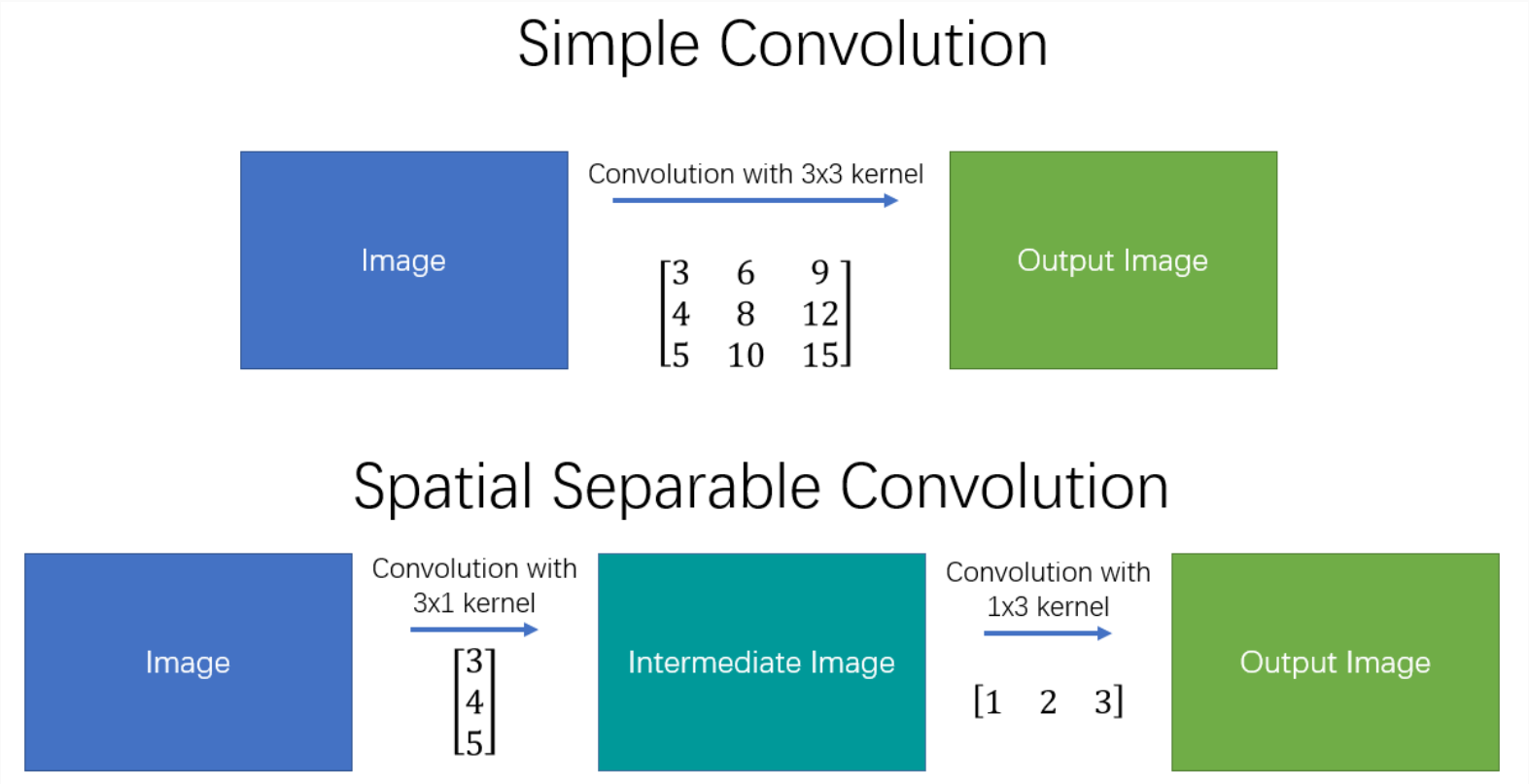


Imagen tomada de `https://towardsdatascience.com/a-basic-introduction-to-separable-convolutions-b99ec3102728`

Convolución en profundidad

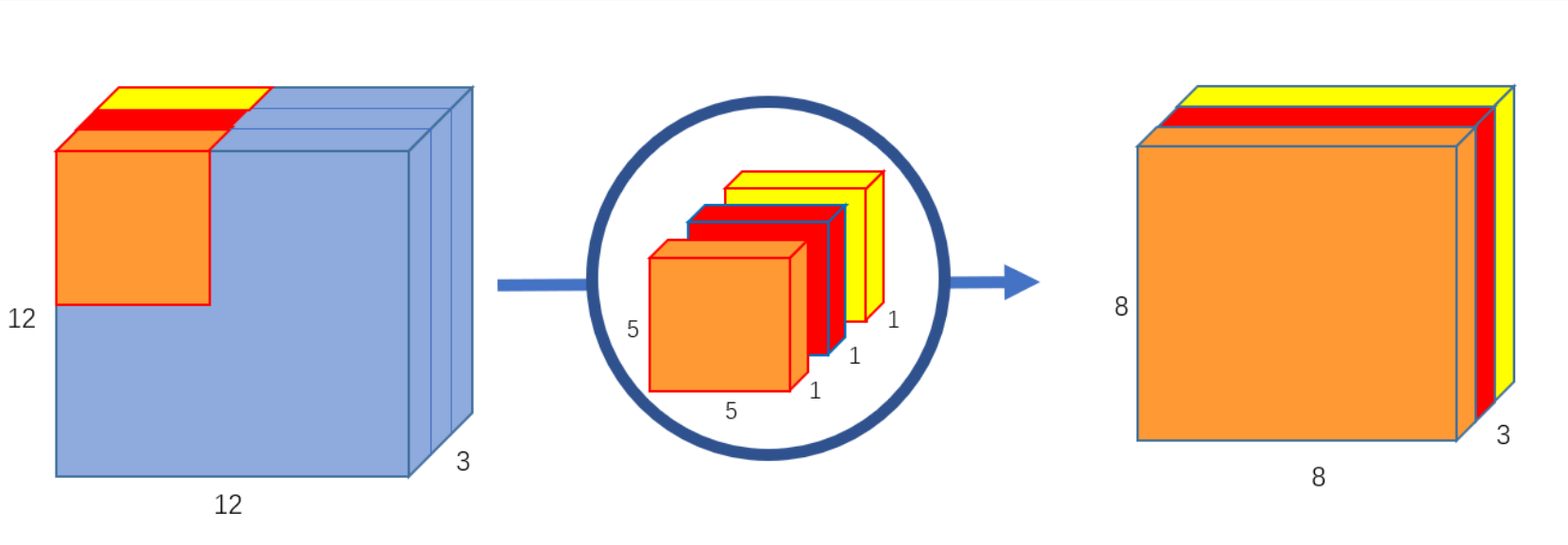


Imagen tomada de https://towardsdatascience.com/a-basic-introduction-to-separable-convolutions-b99ec3102728

Convolución puntual (de 1×1)

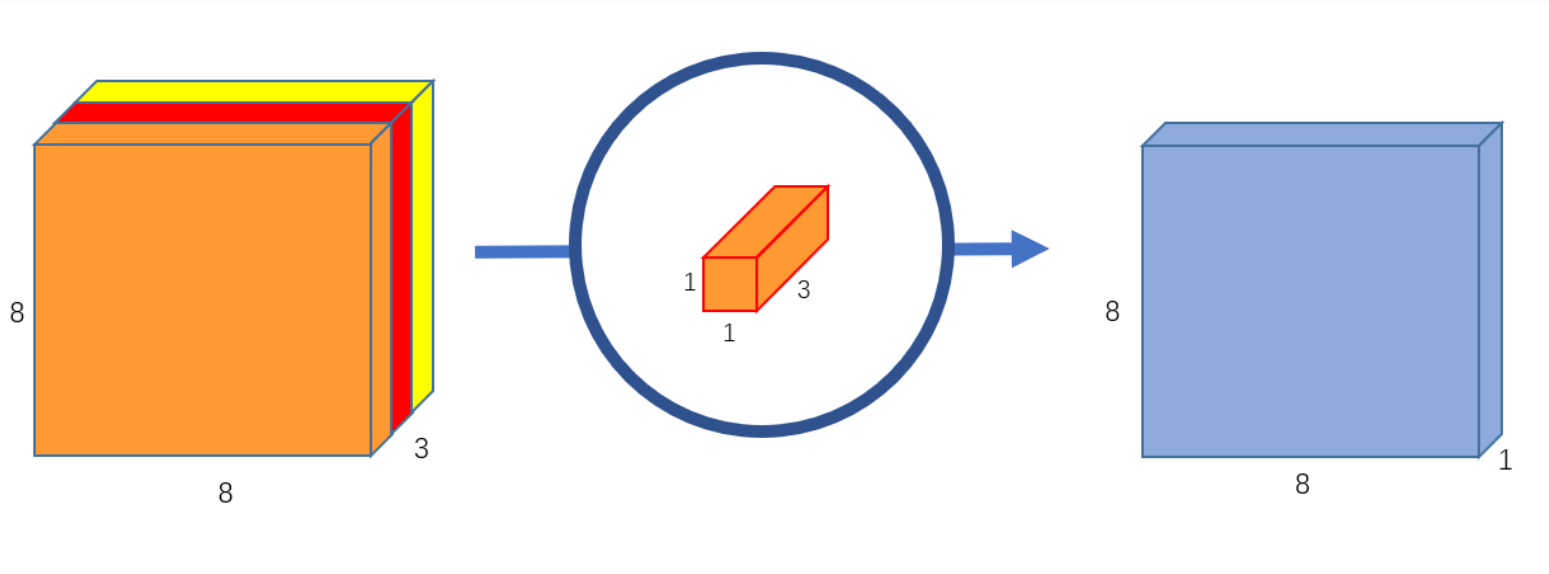


Imagen tomada de `https://towardsdatascience.com/a-basic-introduction-to-separable-convolutions-b99ec3102728`

Convolución separable en profundidad

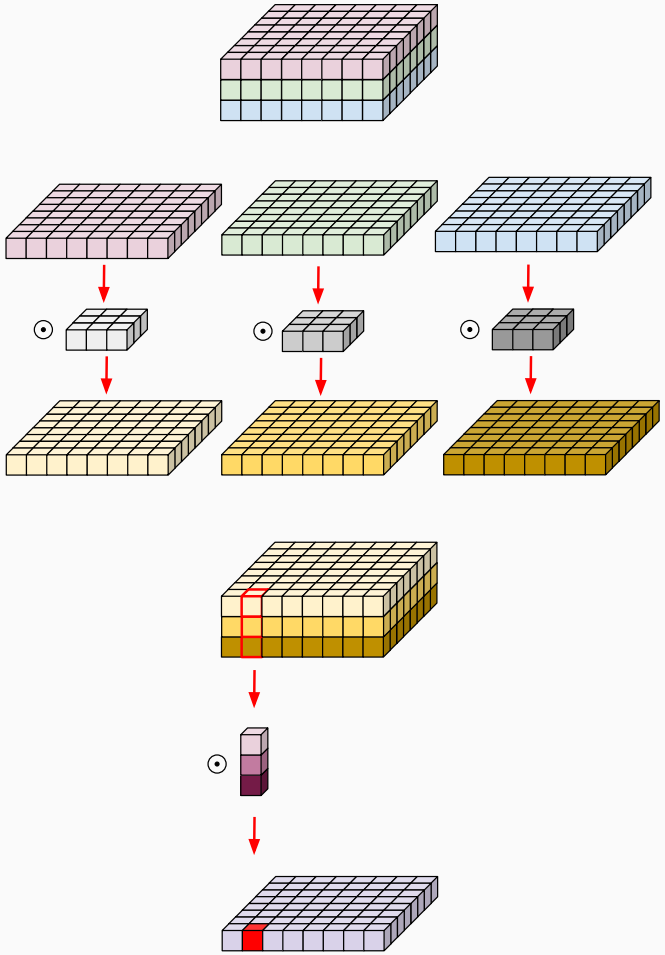


Imagen tomada de https://eli.thegreenplace.net/2018/depthwise-separable-convolutions-for-machine-learning/

Convolución dilatada

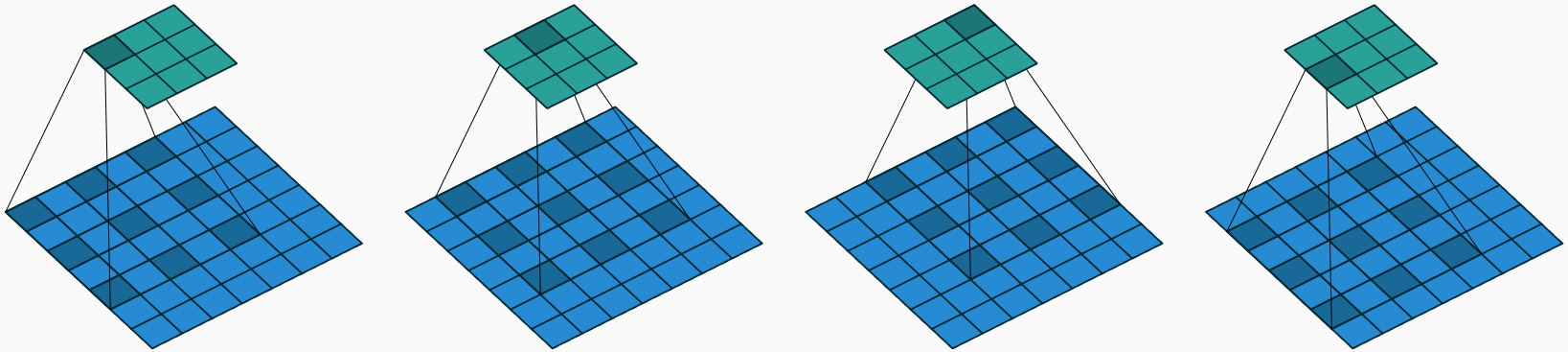


Imagen tomada de Dumoulin and Visin. *A guide to convolution arithmetic for deep learning*, 2018.

Convolución transpuesta

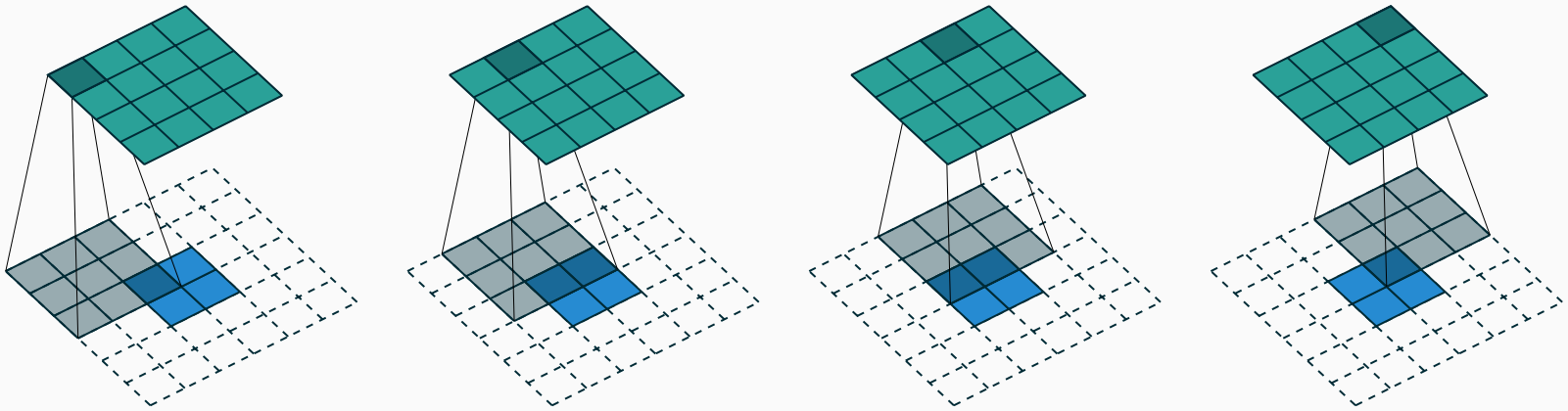


Imagen tomada de Dumoulin and Visin. *A guide to convolution arithmetic for deep learning*, 2018.

Visualizando respuestas

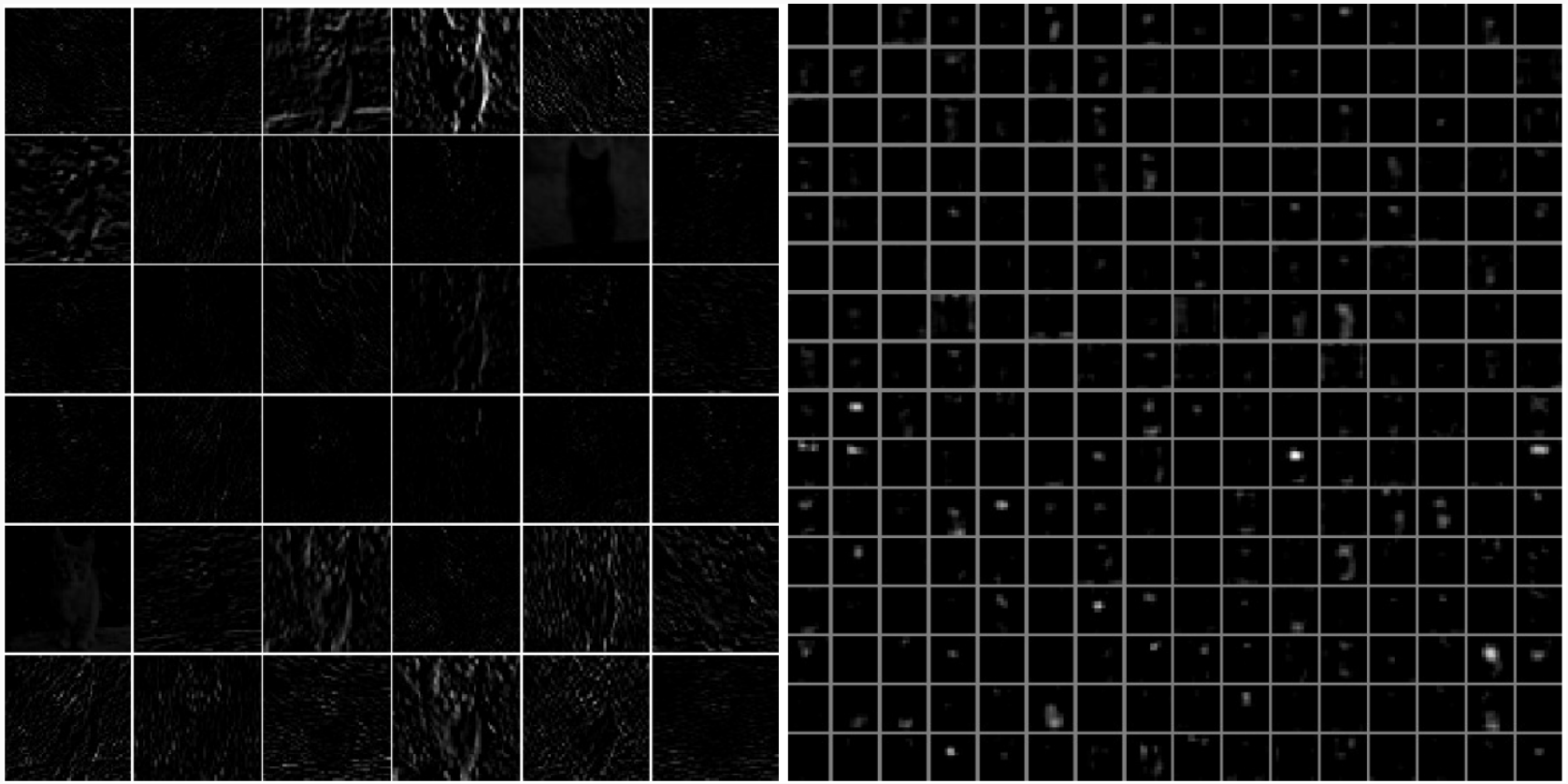


Imagen tomada de `http://cs231n.github.io/understanding-cnn/`

Visualizando pesos

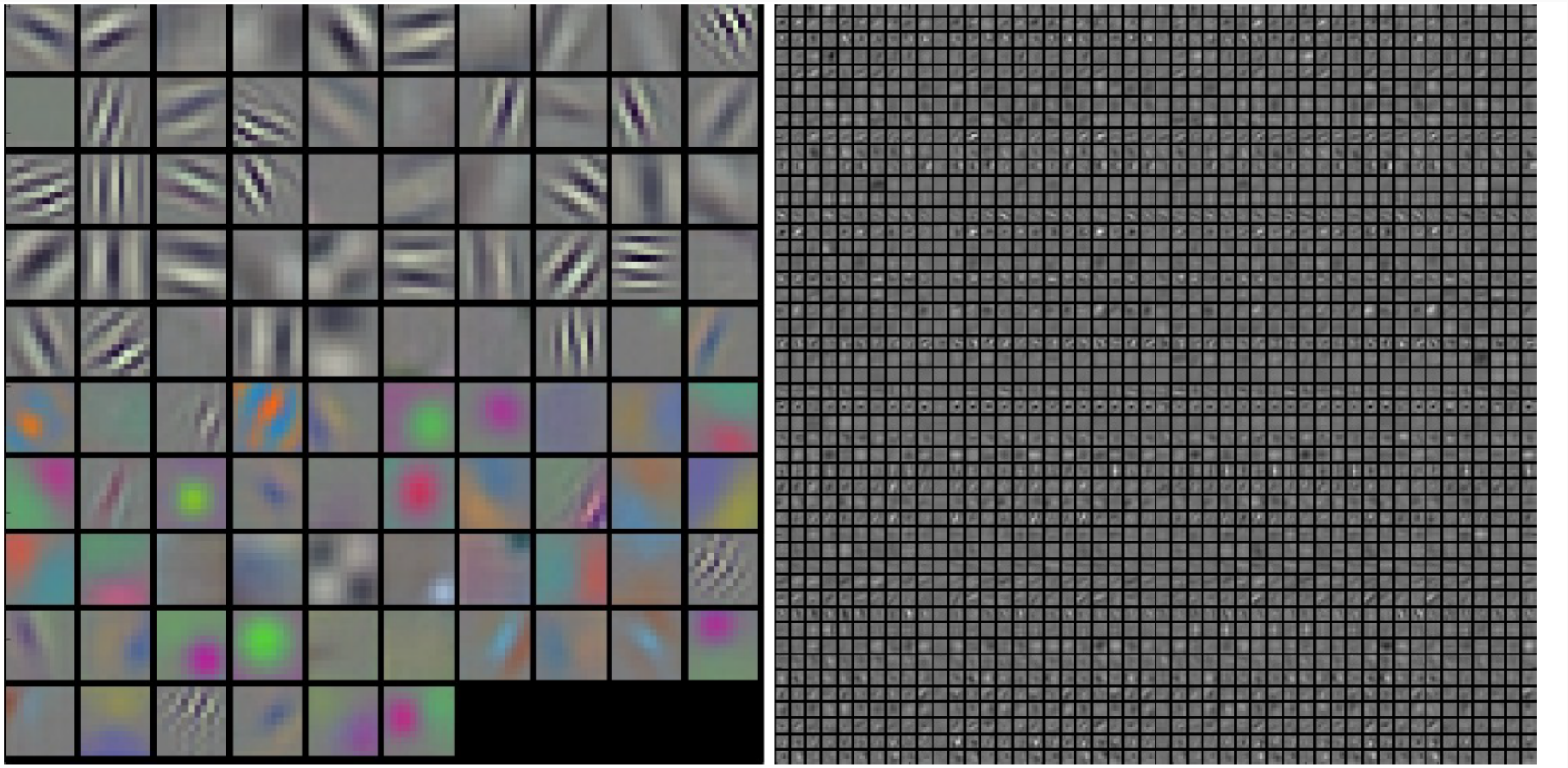


Imagen tomada de `http://cs231n.github.io/understanding-cnn/`

Visualizando imágenes con respuestas máximas

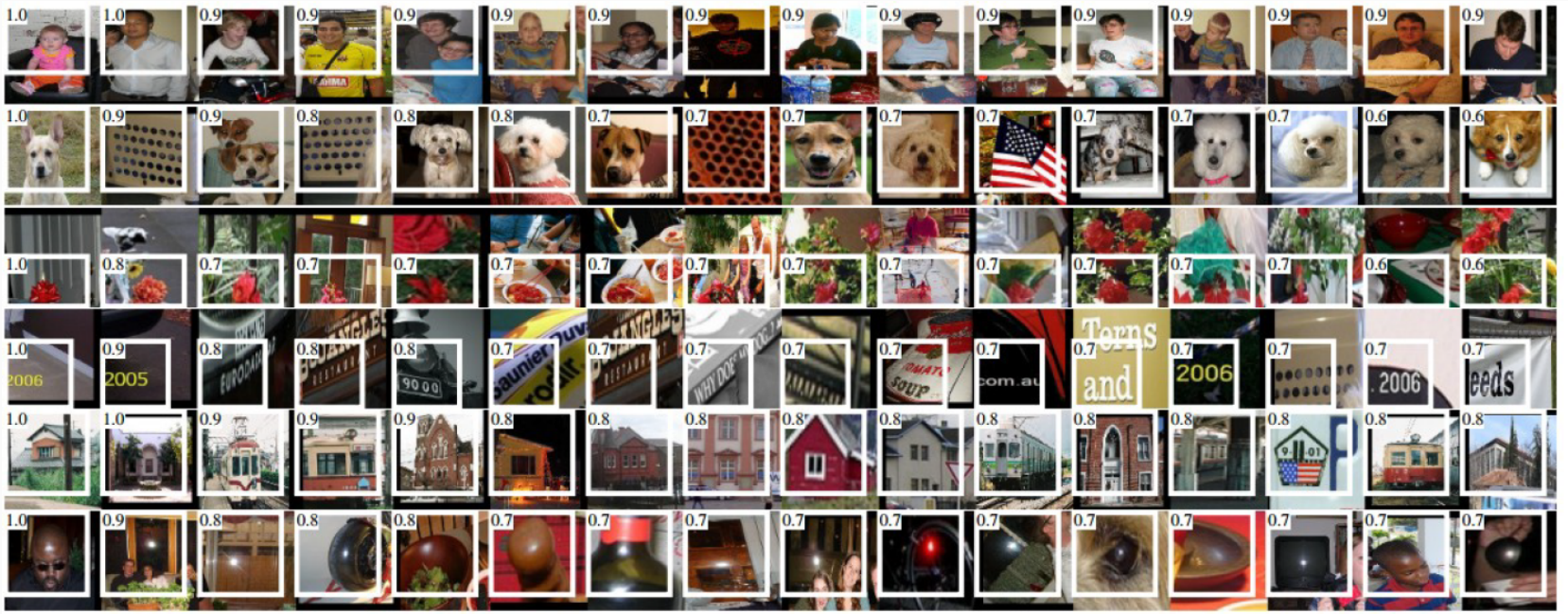


Imagen tomada de `http://cs231n.github.io/understanding-cnn/`

Visualizando obstrucciones

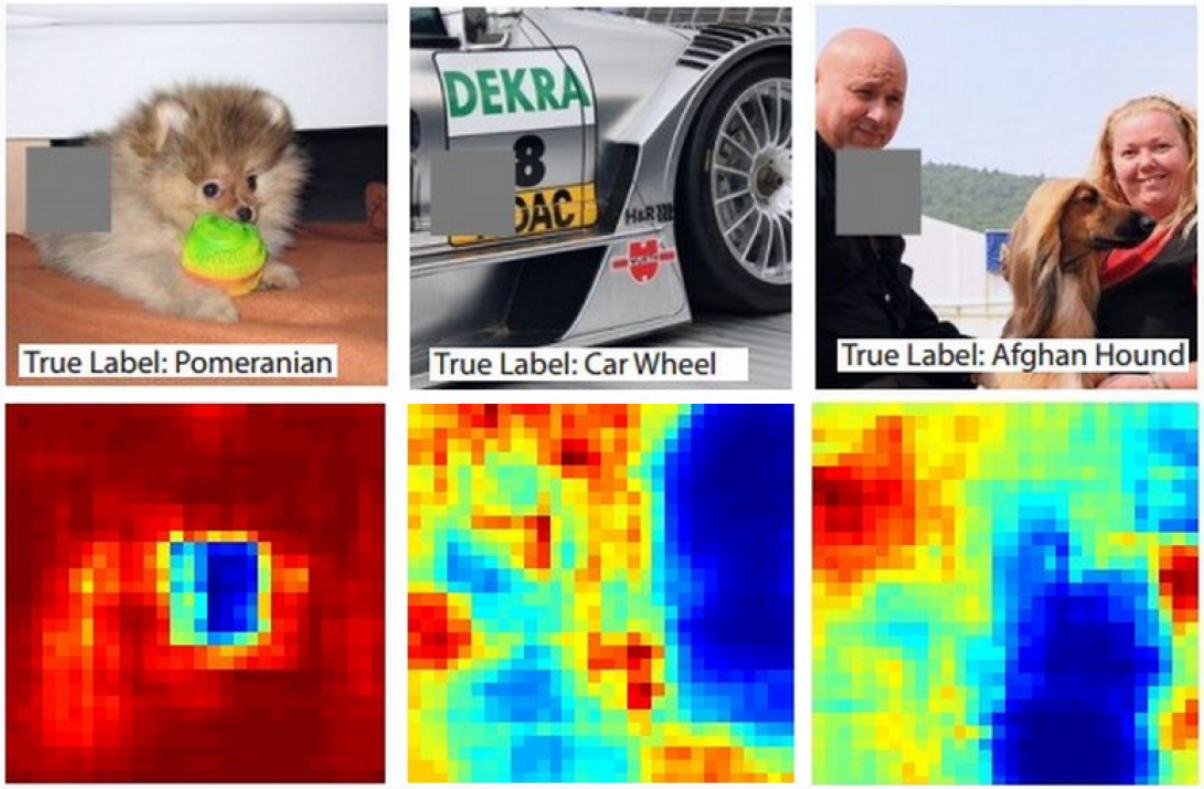


Imagen tomada de `http://cs231n.github.io/understanding-cnn/`

Convolución transpuesta

- Método para visualizar respuestas de diferentes capas en redes neuronales convolucionales

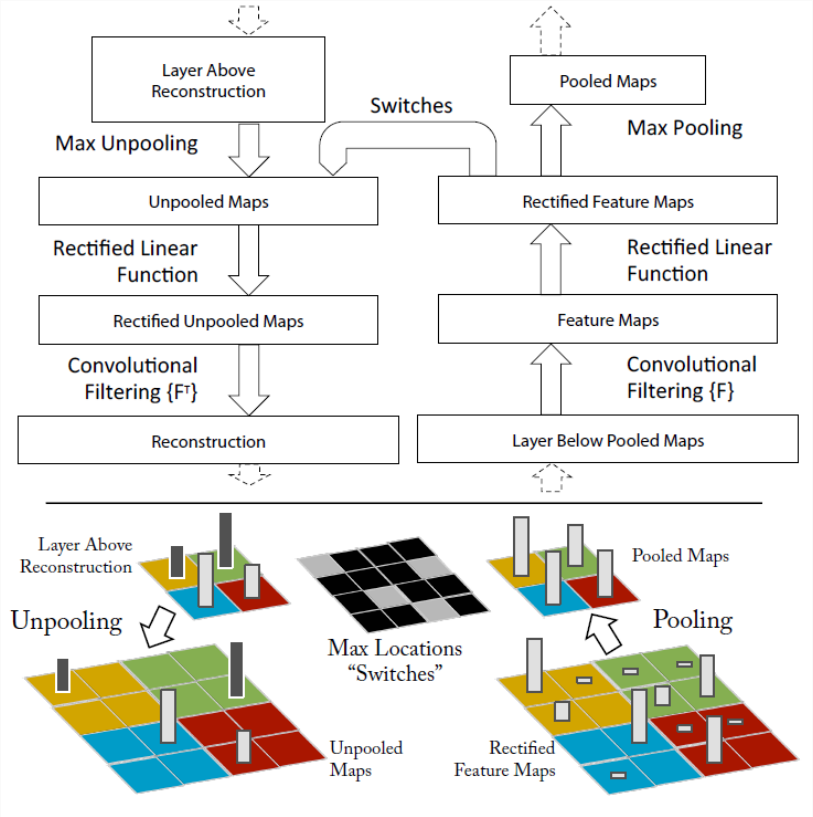


Imagen tomada de Zeiler and Fergus. *Visualizing and Understanding Convolutional Networks*, 2014

DeepDream de Google: diferentes clases

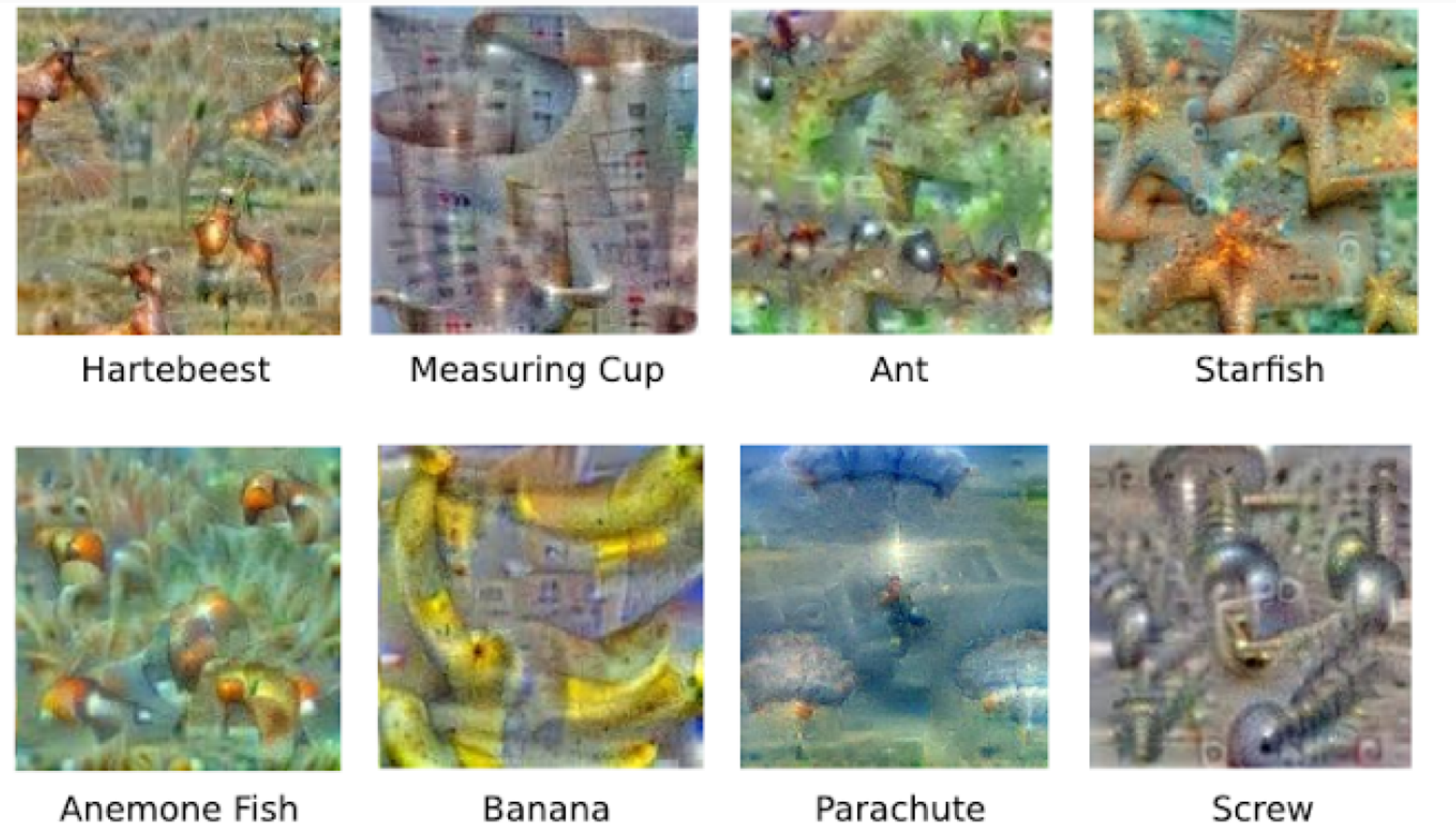


Imagen tomada de `https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html`

DeepDream de Google: características de bajo nivel

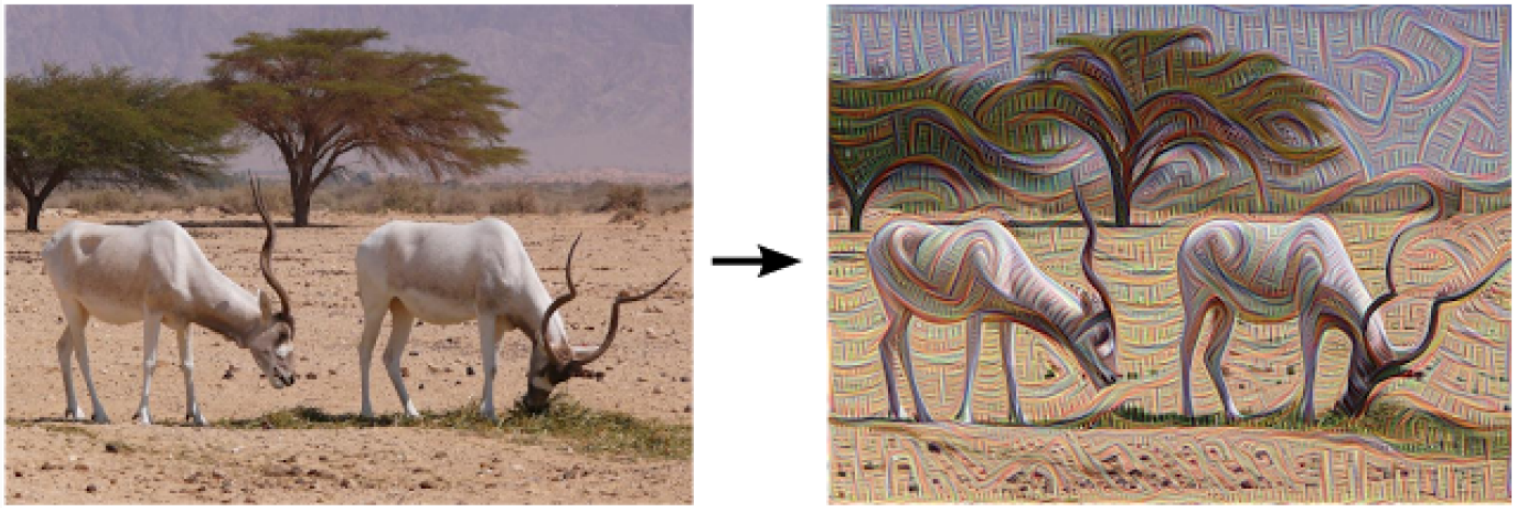


Imagen tomada de https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html

DeepDream de Google: características de alto nivel

Imagen tomada de `https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html`

DeepDream de Google: influencia de imagen original

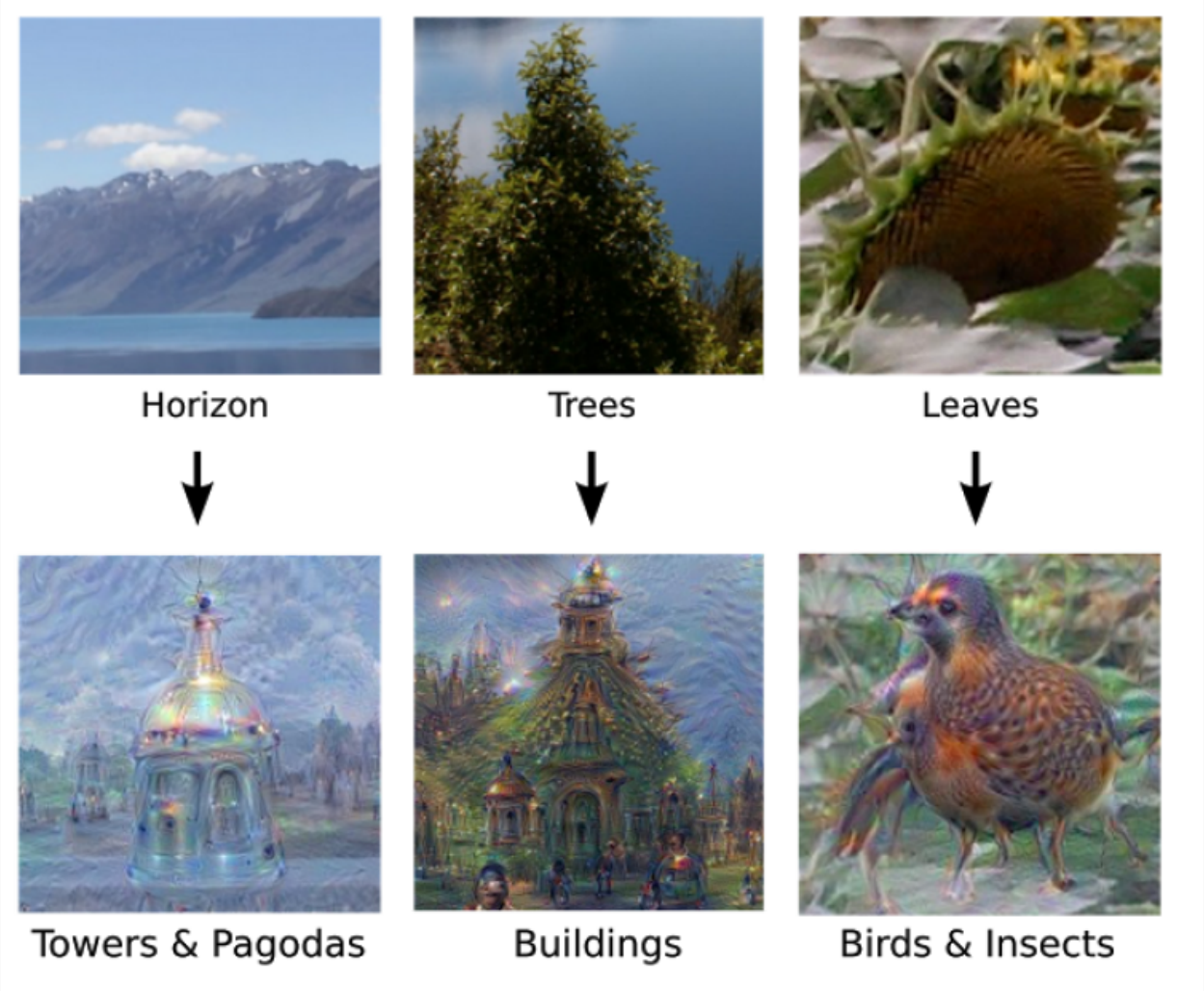


Imagen tomada de `https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html`

DeepDream de Google: de forma aleatoria e iterativa

Imagen tomada de `https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html`