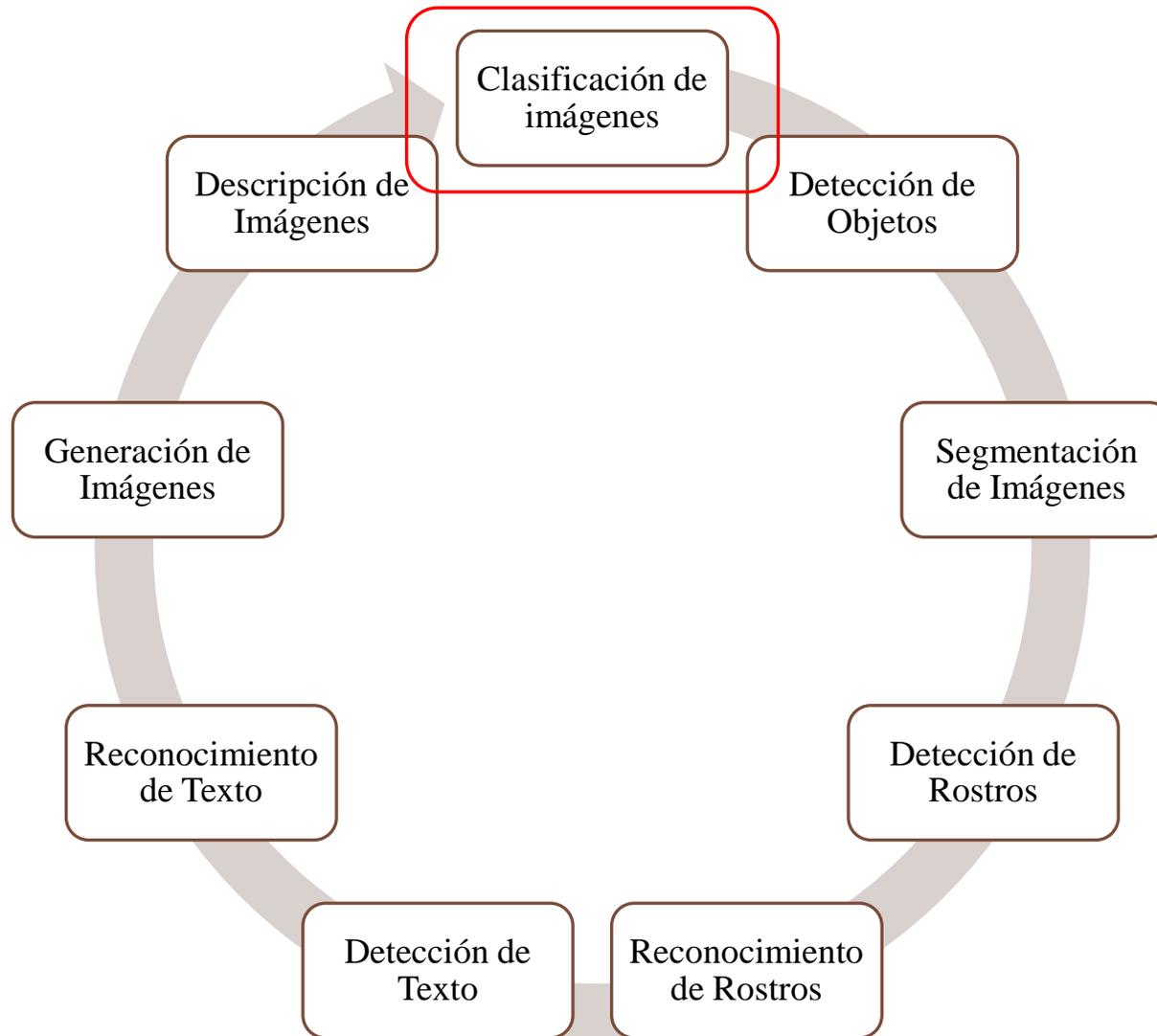


Redes Neuronales Convolucionales

EL 7007 INTRODUCCIÓN AL PROCESAMIENTO DIGITAL DE IMÁGENES

PROFESOR: CLAUDIO PÉREZ

AUXILIAR: JORGE ZAMBRANO

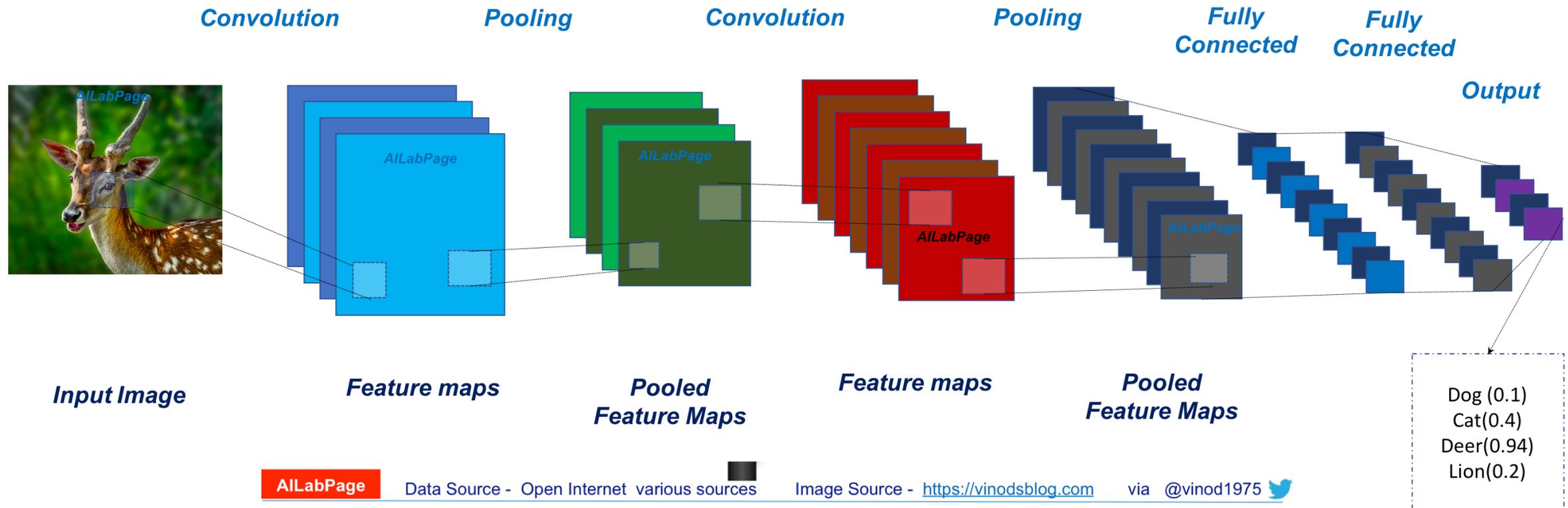


Aplicaciones

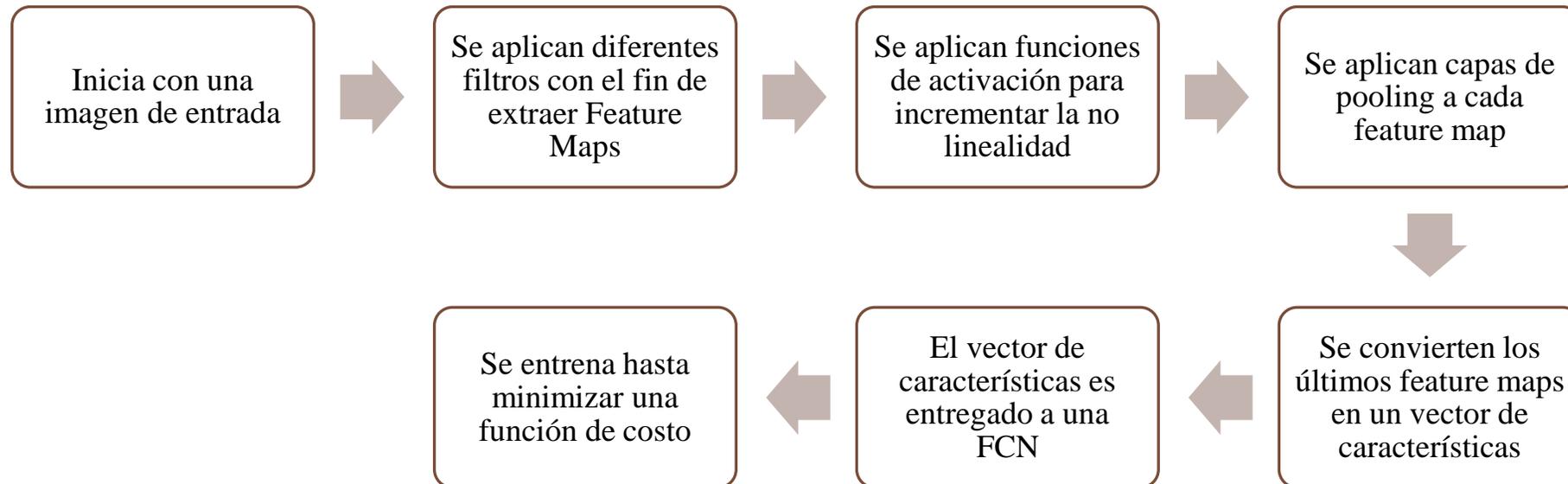
CLASIFICACIÓN DE IMÁGENES

Introducción

Convolution Neural Network



Típicamente una CNN:



Input -> Convolution -> ReLU -> Convolution -> ReLU -> Pooling -> ReLU -> Convolution -> ReLU -> Pooling -> Fully Connected

Capas de Convolución

Feature Maps

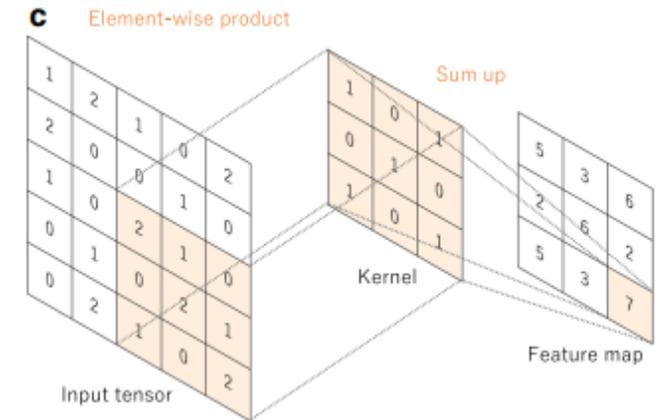
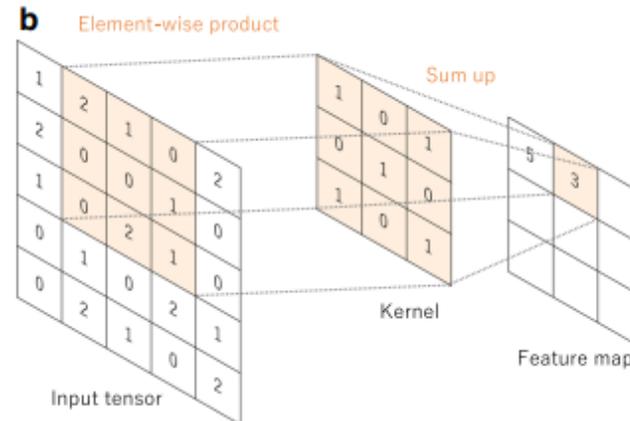
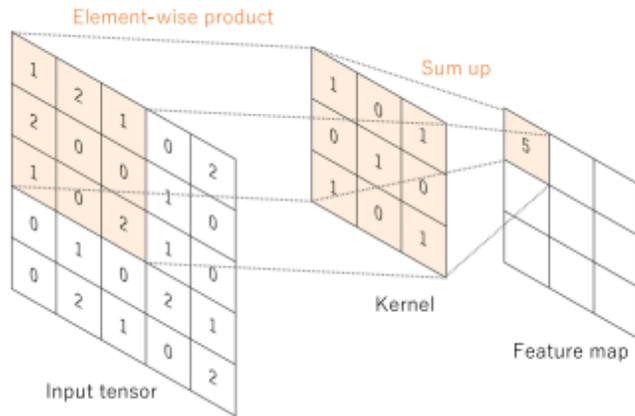
Funciones de activación

Capas de Pooling

Capa Fully Connected (Dense)

Batch Normalization

Elementos y Conceptos típicos de una CNN

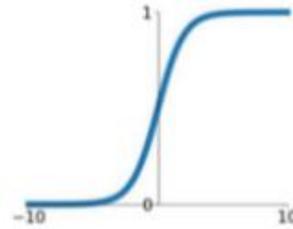


Capas de Convolución y Feature Maps

Funciones de Activación:

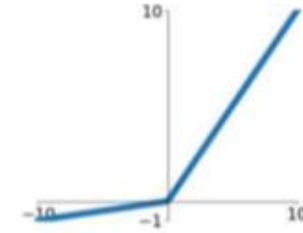
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



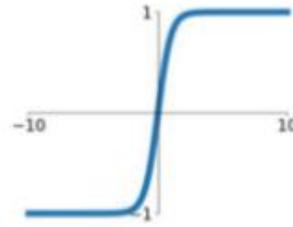
Leaky ReLU

$$\max(0.1x, x)$$



tanh

$$\tanh(x)$$

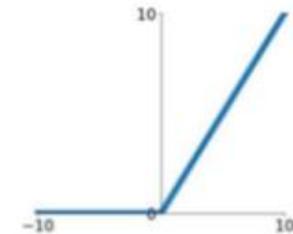


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

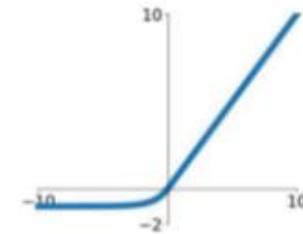
ReLU

$$\max(0, x)$$



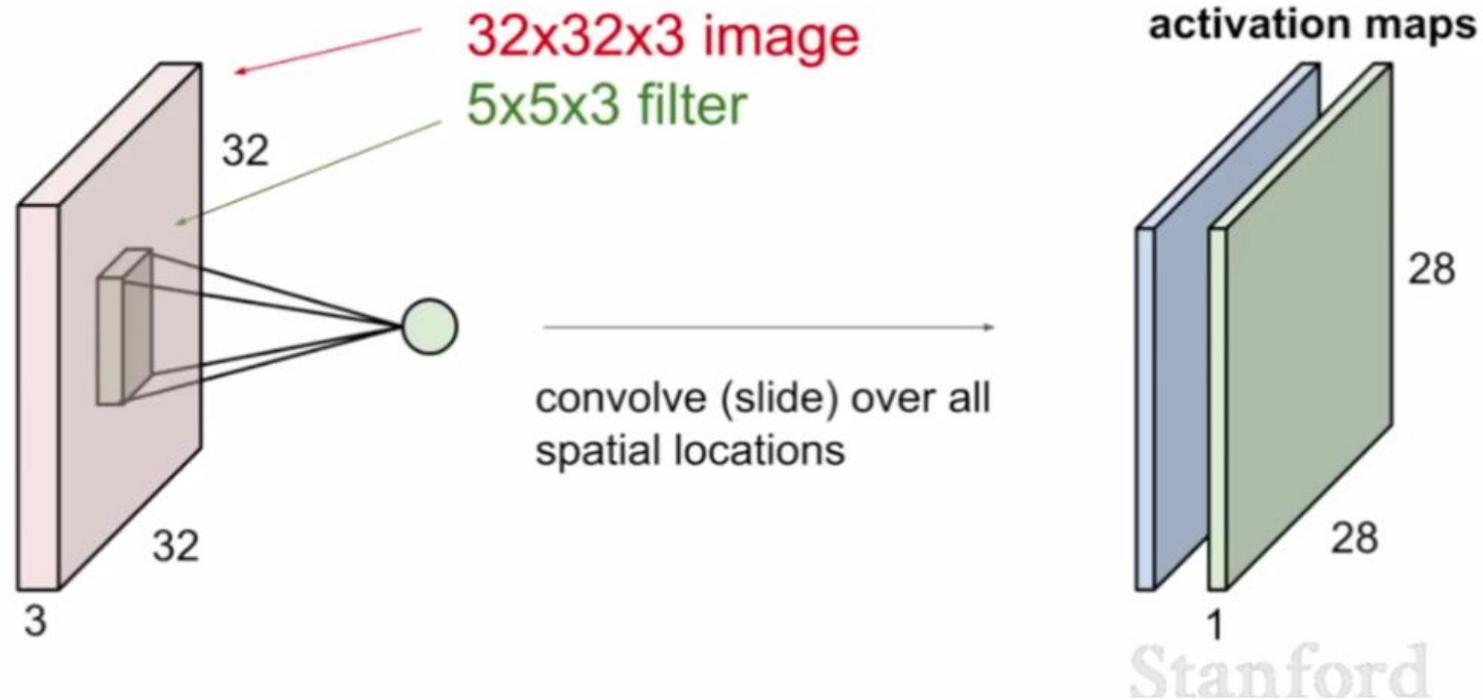
ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



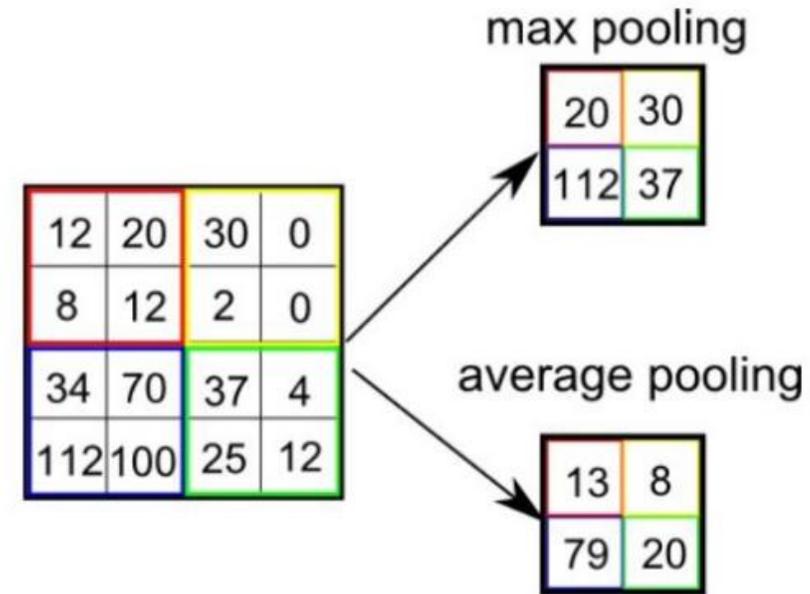
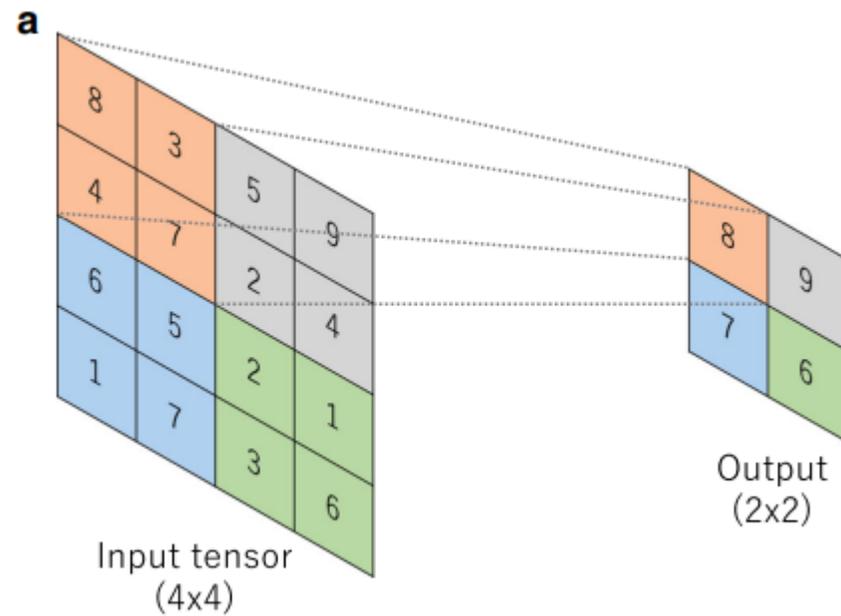
Activation Functions

Funciones de Activación:



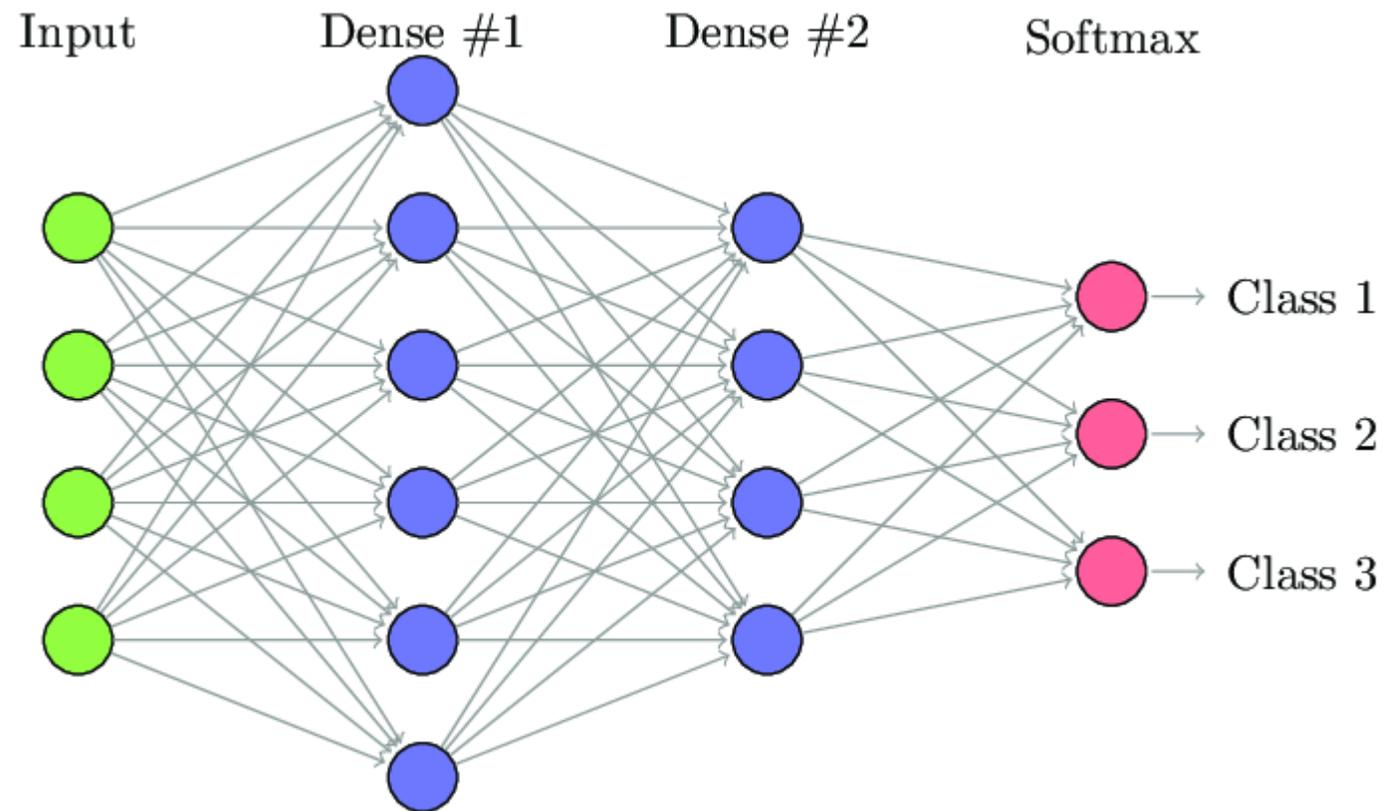
Source: http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture6.pdf

Capas de Pooling:



Types of Pooling

Capas Fully Connected:



Batch Normalization:

$$\text{BN}(\mathbf{x}) = \gamma \odot \frac{\mathbf{x} - \hat{\boldsymbol{\mu}}_{\mathcal{B}}}{\hat{\boldsymbol{\sigma}}_{\mathcal{B}}} + \beta.$$

$$\hat{\boldsymbol{\mu}}_{\mathcal{B}} = \frac{1}{|\mathcal{B}|} \sum_{\mathbf{x} \in \mathcal{B}} \mathbf{x},$$
$$\hat{\boldsymbol{\sigma}}_{\mathcal{B}}^2 = \frac{1}{|\mathcal{B}|} \sum_{\mathbf{x} \in \mathcal{B}} (\mathbf{x} - \hat{\boldsymbol{\mu}}_{\mathcal{B}})^2 + \epsilon.$$

... Depende del Batch y no del Dataset.

Parámetros: Gamma y Beta

Para capas densas:

Se aplica a la salida y después una FA

Para capas convolucionales:

Se aplica a cada canal

Clasificación de
Imágenes.
Redes más
conocidas...

AlexNet [1]

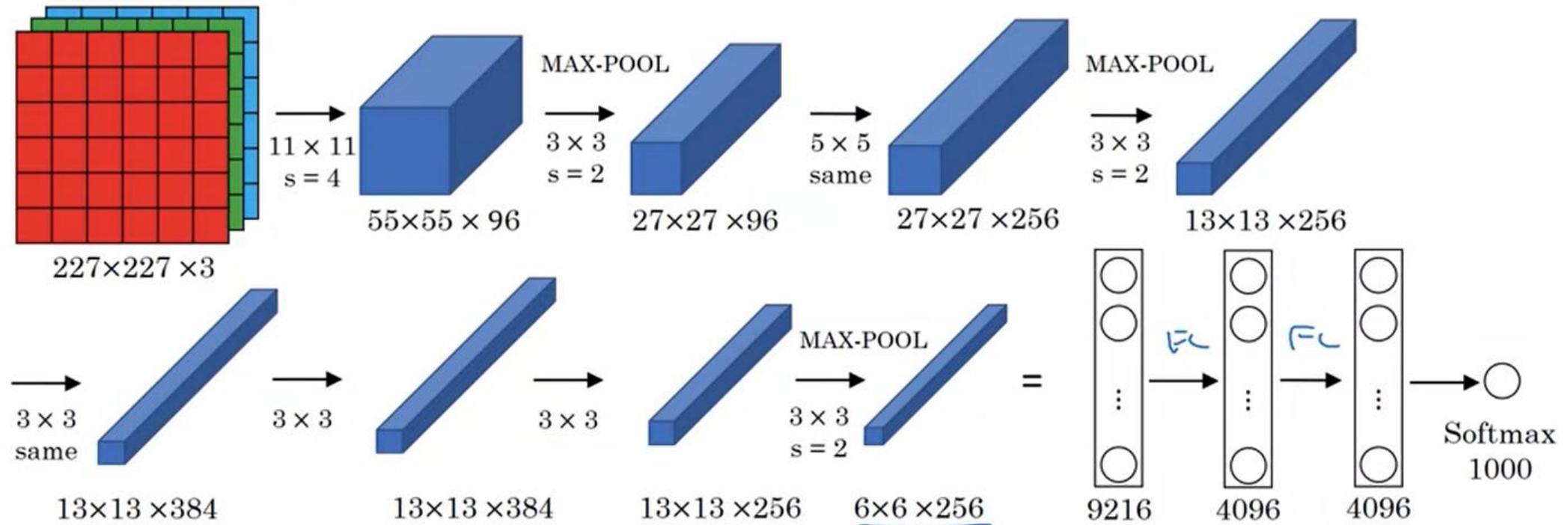
VGG [2]

GoogLeNet [3]

ResNet [4]

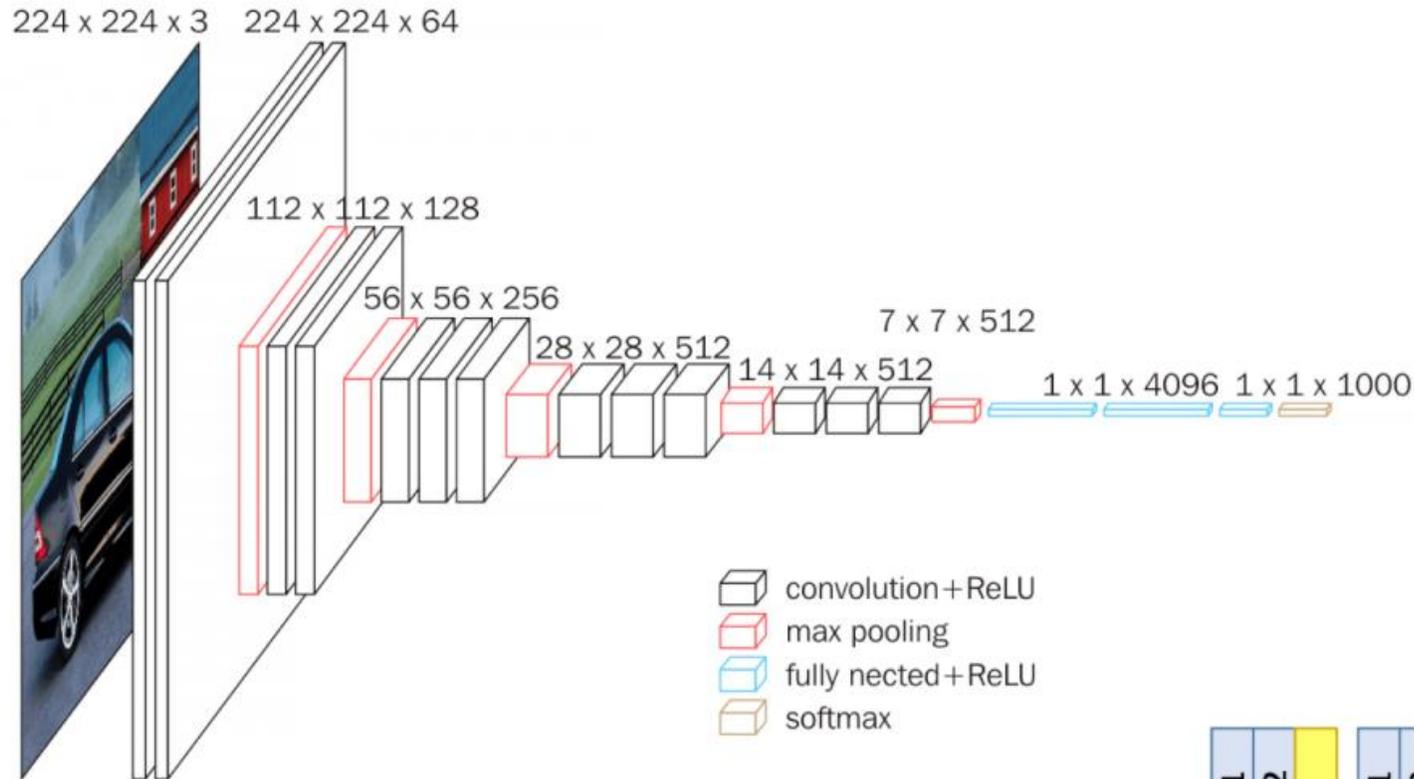
DenseNet [5]

AlexNet (2012)



AlexNet won the 2012 ImageNet competition with a top-5 error rate of 15.3%, compared to the second place top-5 error rate of 26.2%.

VGG (2014)

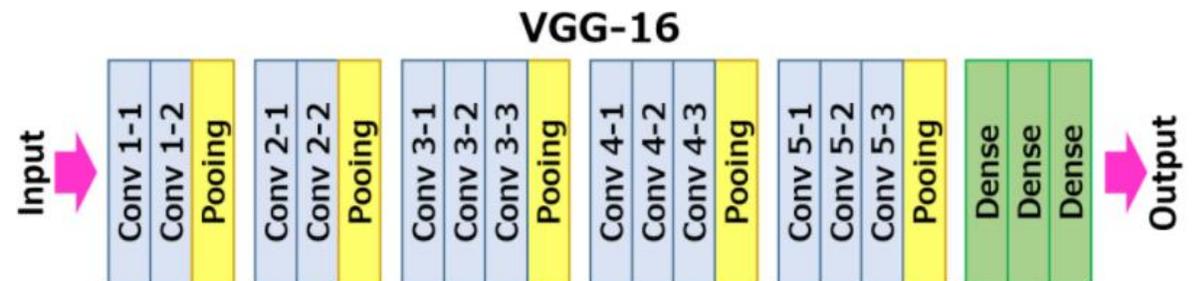


Arquitectura de VGG 16

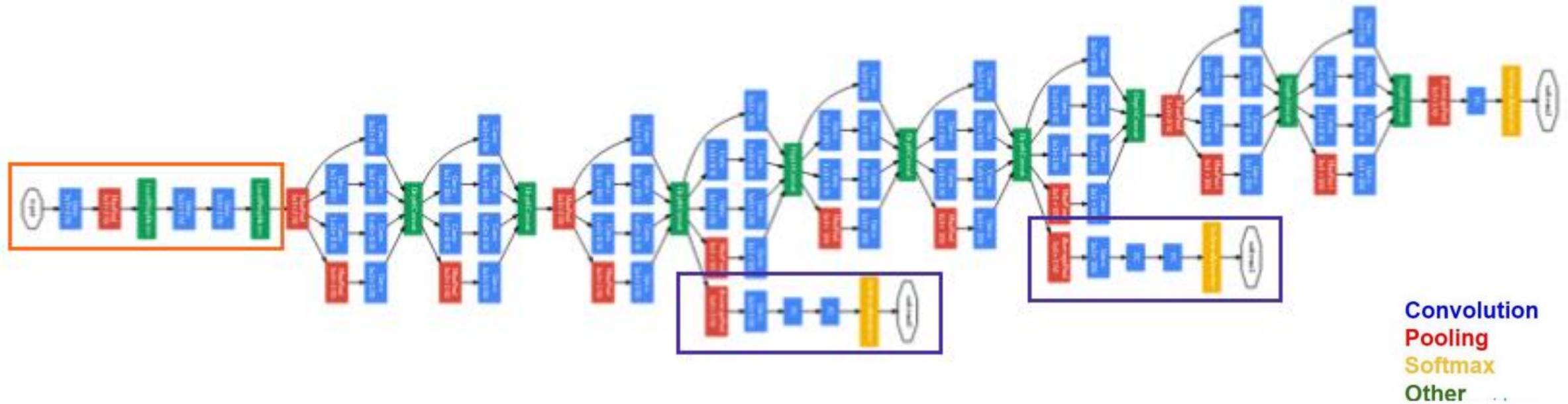
Filtros más pequeños..

Kernel: 3x3

The network achieved 92.7% top-5 test accuracy on the ImageNet dataset.

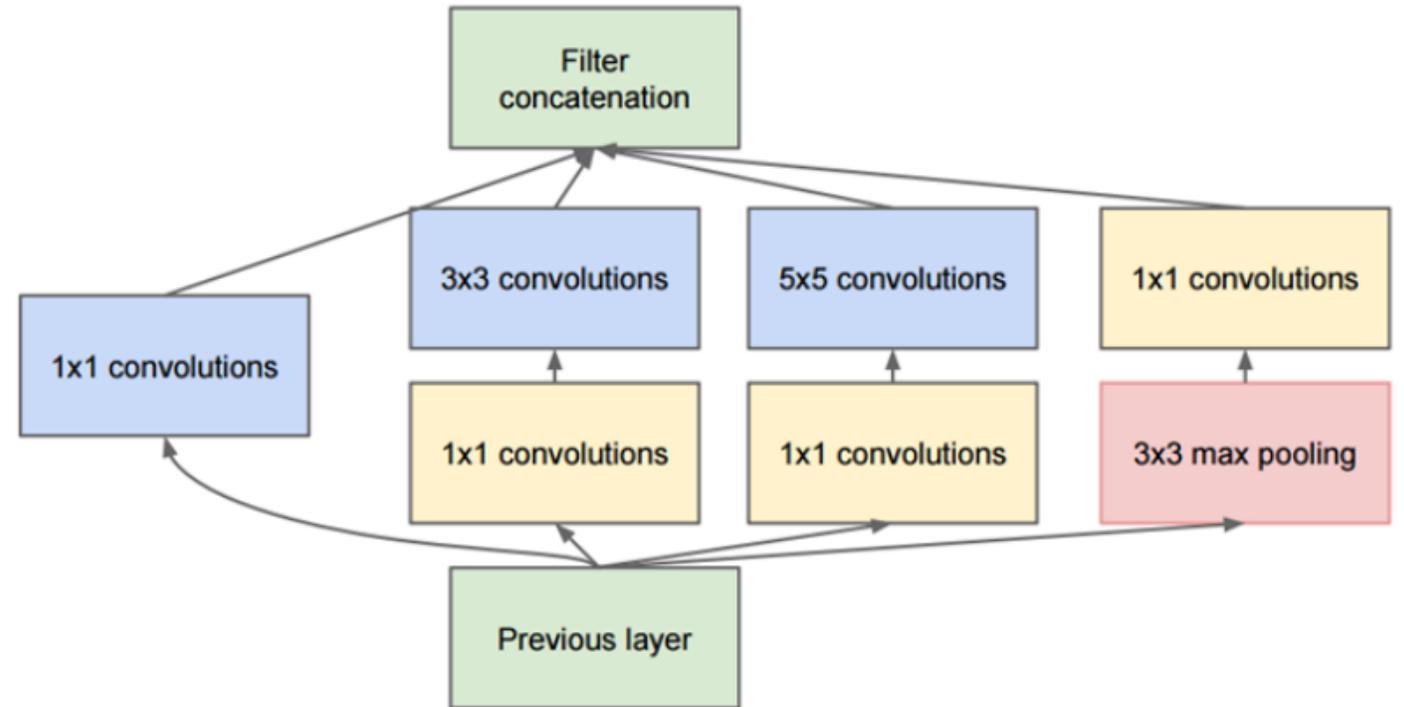
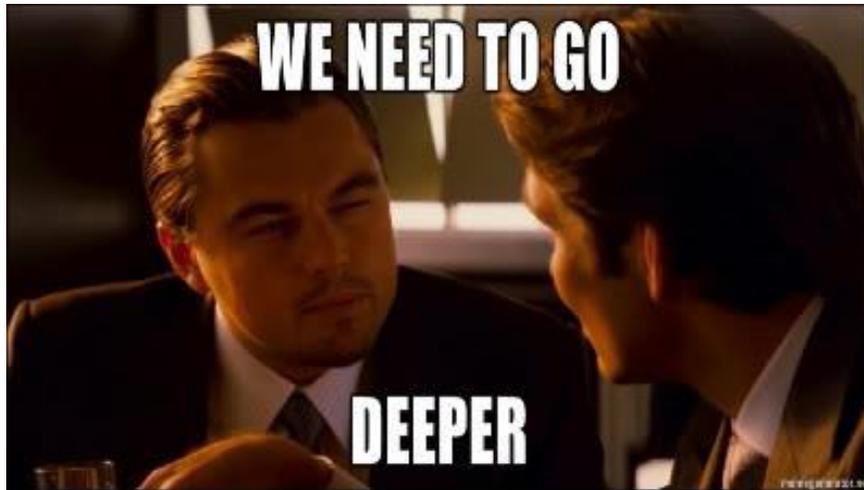


GoogLeNet (2014)



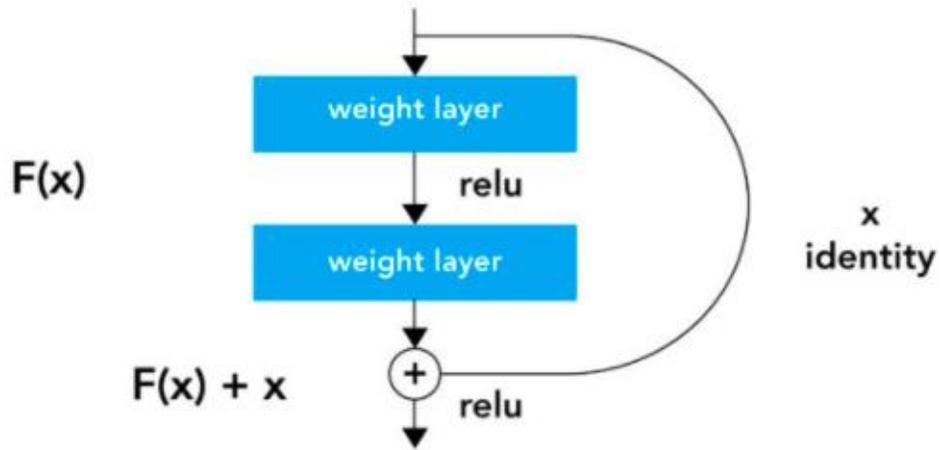
GoogLeNet was the winner at ILSRVRC 2014 taking 1st place in both classification and detection tasks. It has a top-5 error rate of 6.67% in classification task.

GoogLeNet (2014)

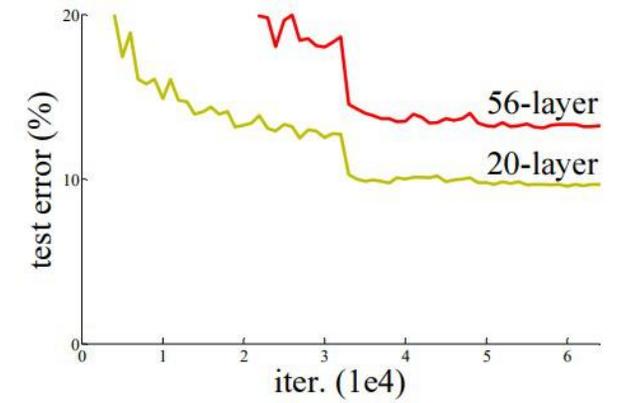
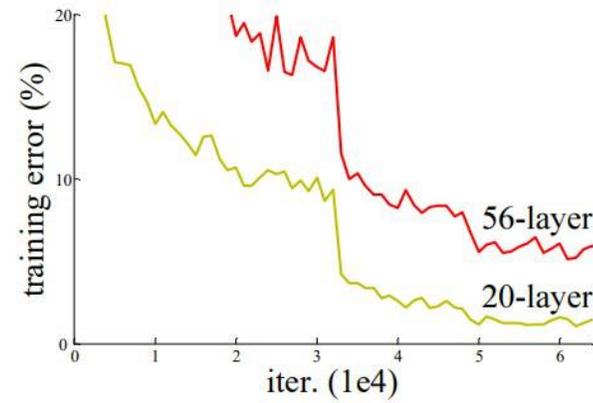


Full Inception module

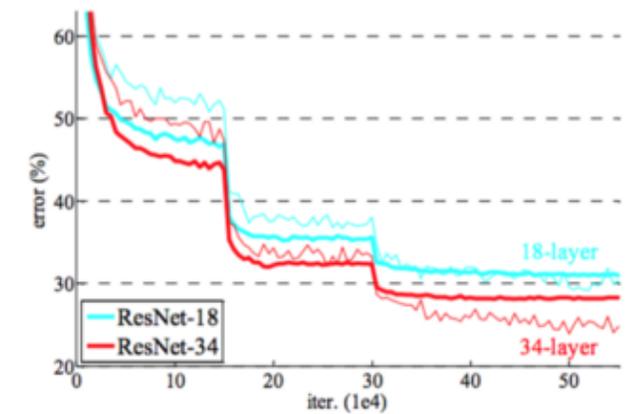
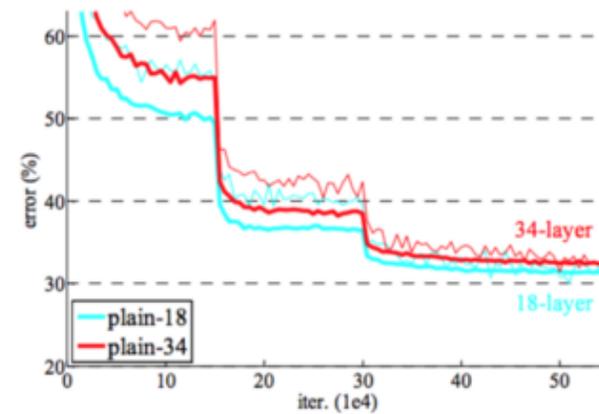
ResNet (2015)



Bloque Residual de ResNet

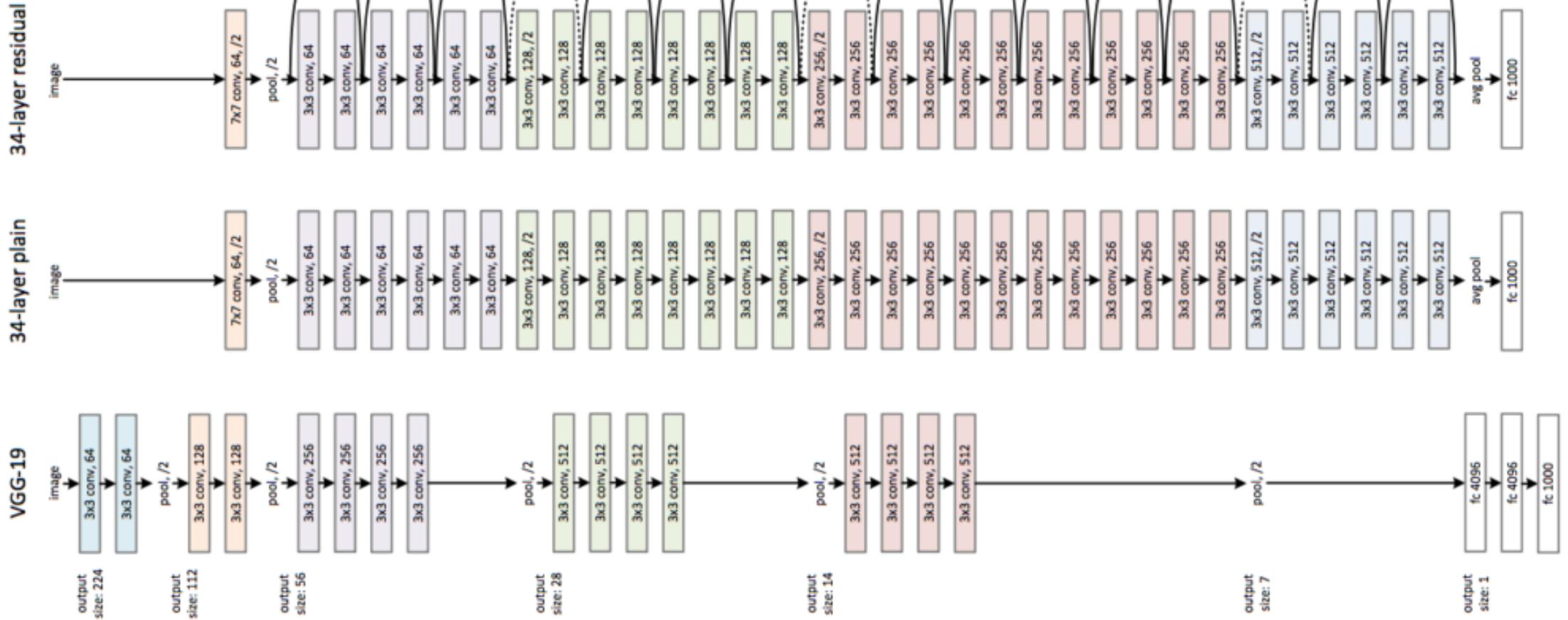


Más Capas .. ¿Más error?

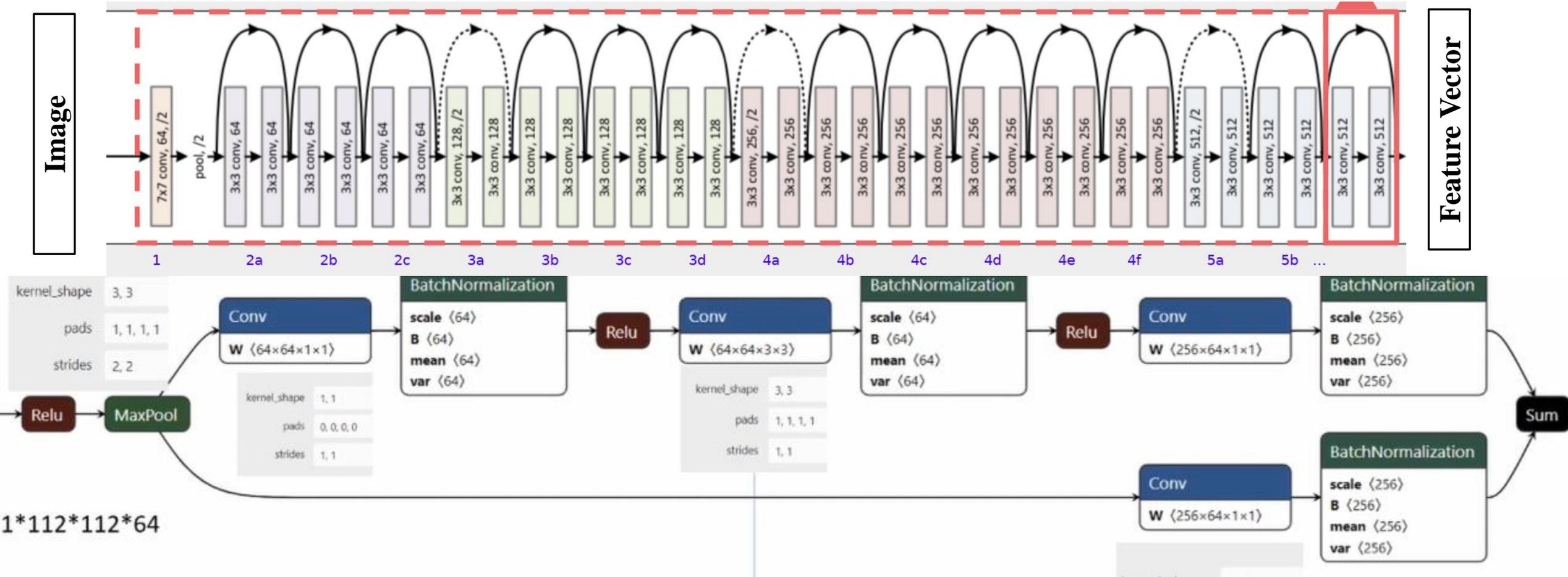


Resultados de ResNet

ResNet (2015)



ResNet (2015)



DenseNet (2017)

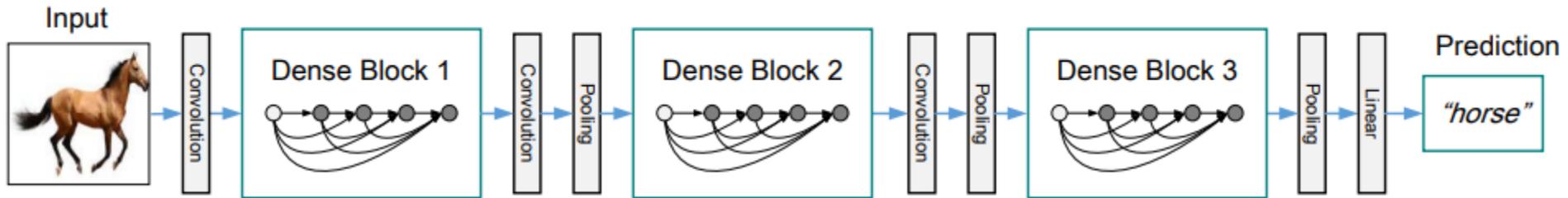


Figure 2: A deep DenseNet with three dense blocks. The layers between two adjacent blocks are referred to as transition layers and change feature-map sizes via convolution and pooling.

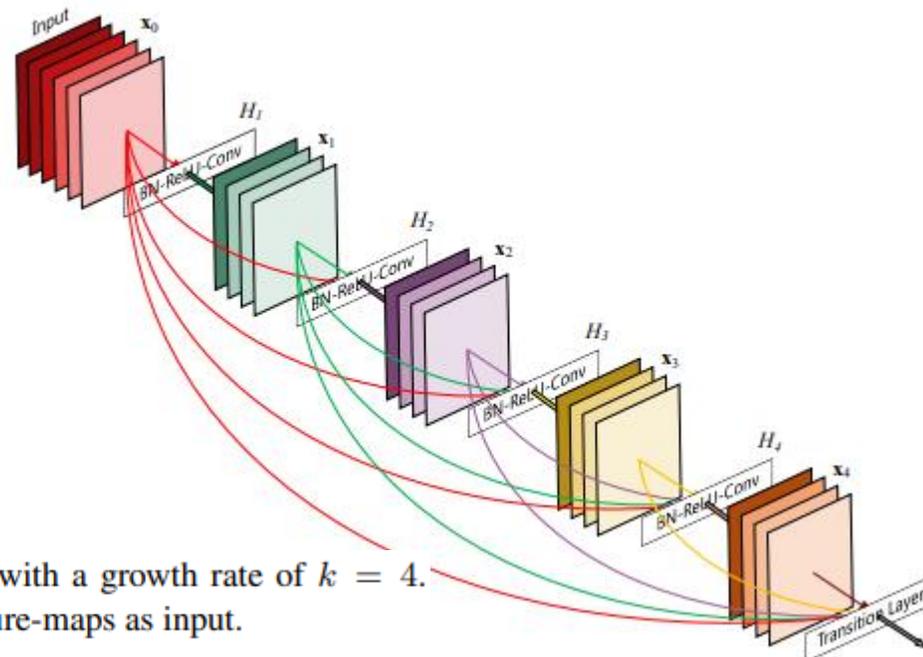
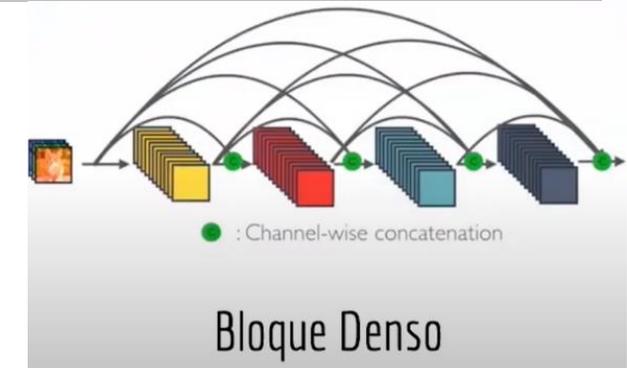
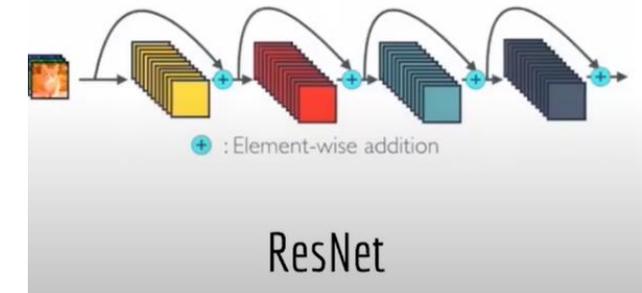
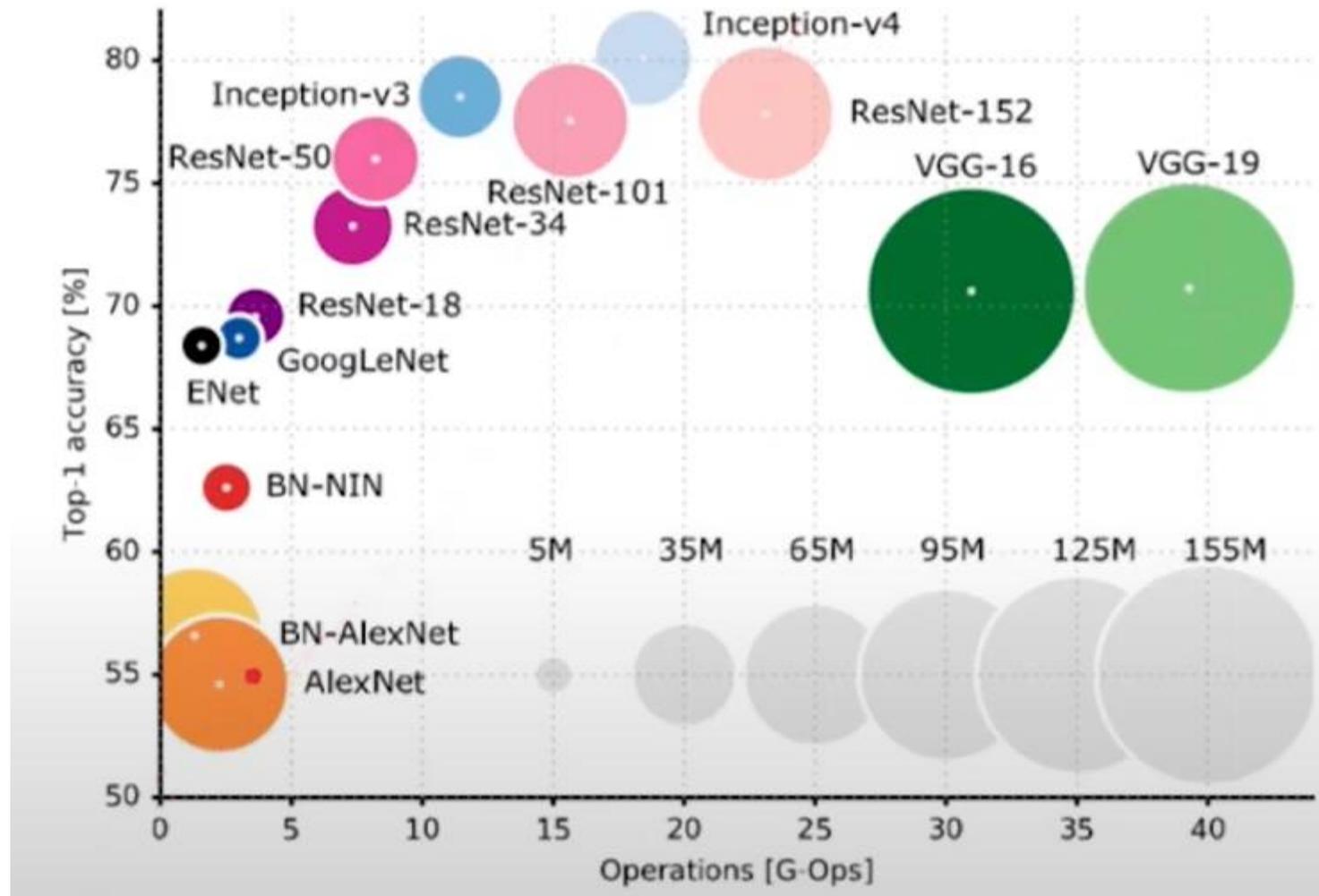


Figure 1: A 5-layer dense block with a growth rate of $k = 4$. Each layer takes all preceding feature-maps as input.



Conclusión



Redes Neuronales Convolucionales

EL 7007 INTRODUCCIÓN AL PROCESAMIENTO DIGITAL DE IMÁGENES

PROFESOR: CLAUDIO PÉREZ

AUXILIAR: JORGE ZAMBRANO