

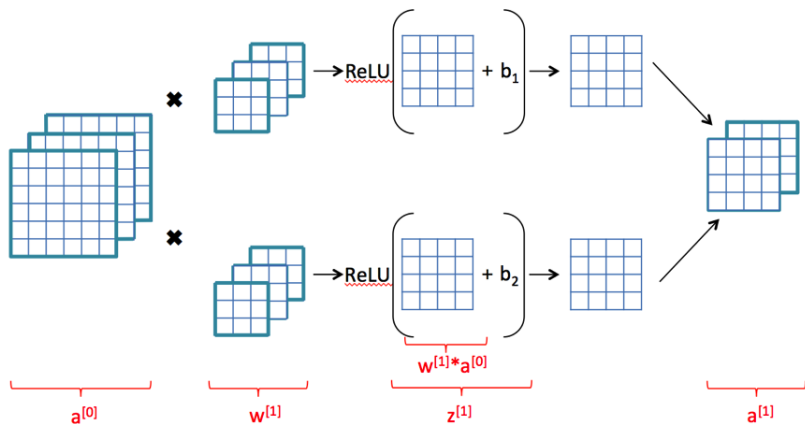
Deep Learning

Vazgen Mikayelyan

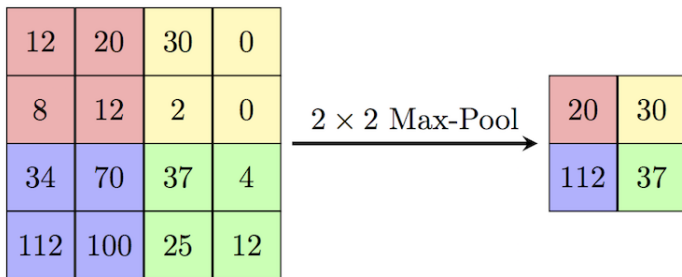
November 17, 2020



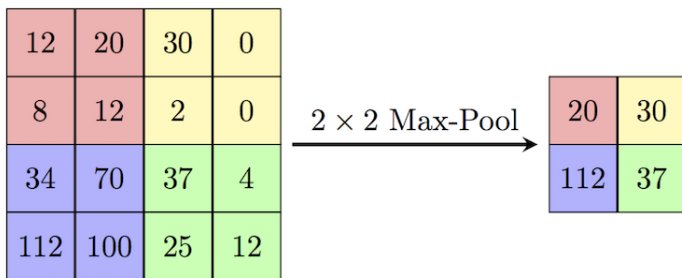
One Layer of CNN



Pooling Layers

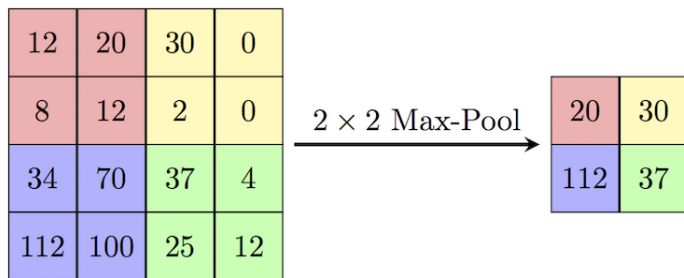


Pooling Layers



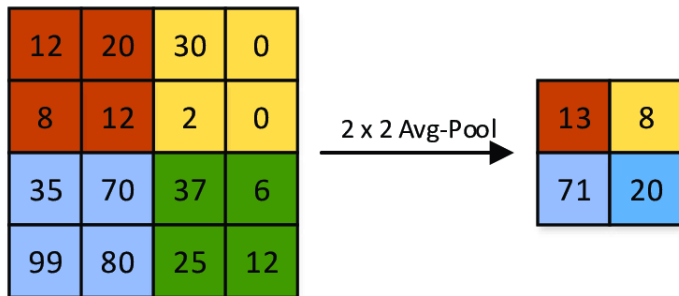
- Number of channels is the same after pooling layer.

Pooling Layers



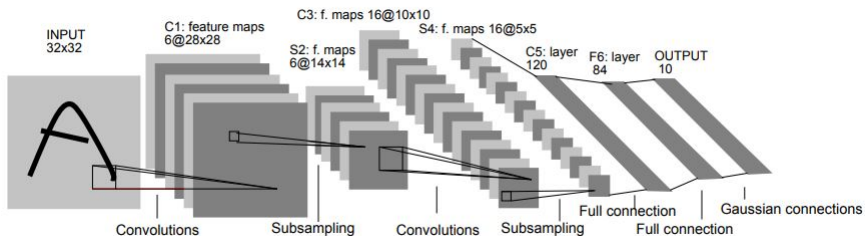
- Number of channels is the same after pooling layer.
- There are not trainable parameters in this layer.

Pooling Layers

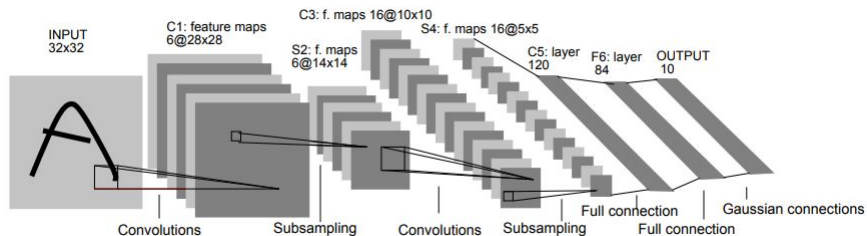


1 Famous CNNs

LeNet-5 (1998)

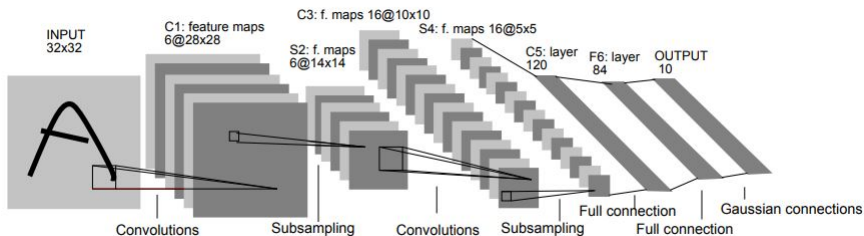


LeNet-5 (1998)



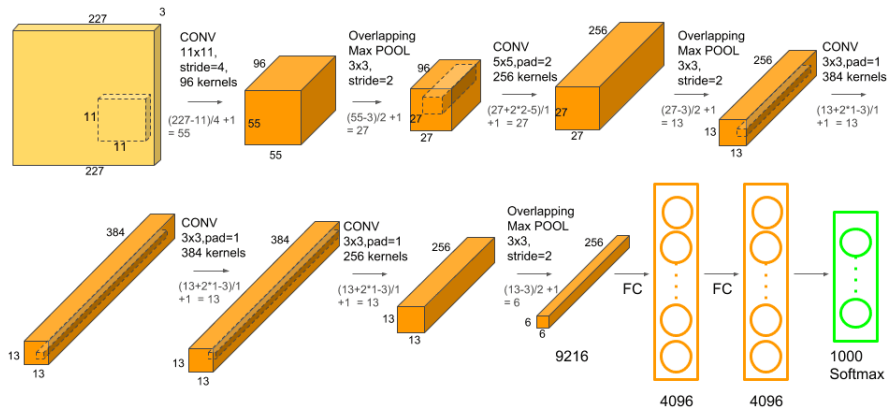
- Activation functions are sigmoids and hyperbolic tangents.

LeNet-5 (1998)



- Activation functions are sigmoids and hyperbolic tangents.
- LeNet-5 has approximately 60k parameters.

AlexNet (2012)



- The ReLU non-linearity is applied to the output of every convolutional and fully-connected layer.

- The ReLU non-linearity is applied to the output of every convolutional and fully-connected layer.
- Local response normalization?

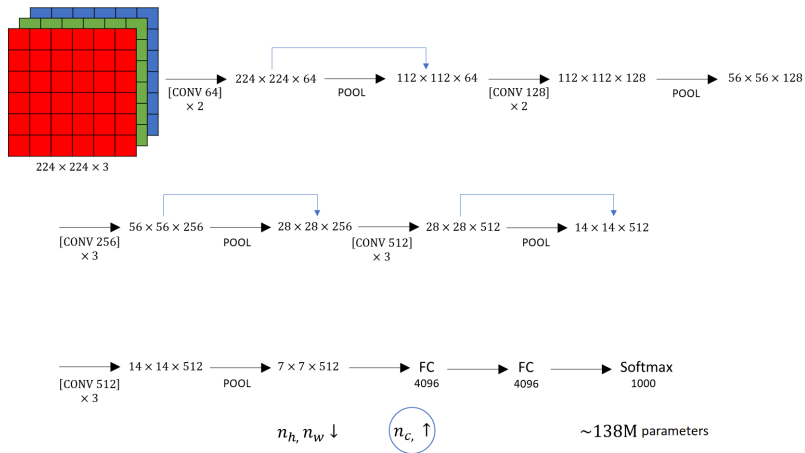
- The ReLU non-linearity is applied to the output of every convolutional and fully-connected layer.
- Local response normalization?
- AlexNet has approximately 60M parameters.

- The ReLU non-linearity is applied to the output of every convolutional and fully-connected layer.
- Local response normalization?
- AlexNet has approximately 60M parameters.
- Accuracies on ImageNet: Top1=63.3%, Top5=84.6%.

VGG-16 (2014)

CONV = 3 x 3 filter, s=1, same

MAX-POOL = 2 x 2, s=2



- Fixed kernel size.

VGG-16 (2014)

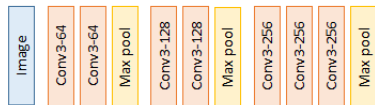
- Fixed kernel size.
- The ReLU non-linearity is applied to the output of every convolutional and fully-connected layer.

VGG-16 (2014)

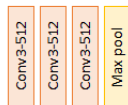
- Fixed kernel size.
- The ReLU non-linearity is applied to the output of every convolutional and fully-connected layer.
- VGG-16 has approximately 138M parameters.

- Fixed kernel size.
- The ReLU non-linearity is applied to the output of every convolutional and fully-connected layer.
- VGG-16 has approximately 138M parameters.
- Accuracies on ImageNet: Top1=74.4%, Top5=91.9%.

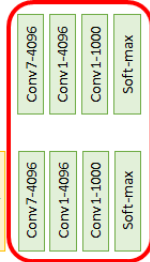
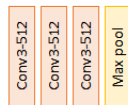
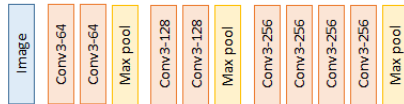
VGG-19 (2014)



VGG-16



VGG-19



- The ReLU non-linearity is applied to the output of every convolutional and fully-connected layer.

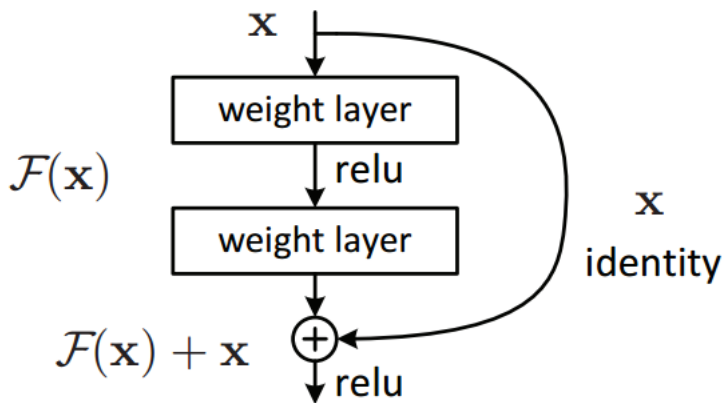
VGG-19 (2014)

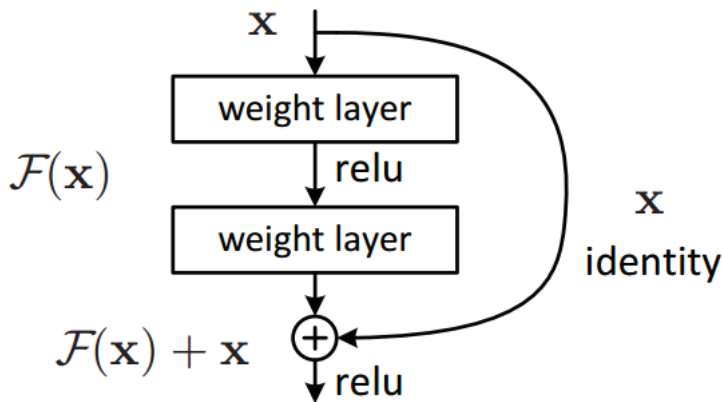
- The ReLU non-linearity is applied to the output of every convolutional and fully-connected layer.
- VGG-19 has approximately 144M parameters.

VGG-19 (2014)

- The ReLU non-linearity is applied to the output of every convolutional and fully-connected layer.
- VGG-19 has approximately 144M parameters.
- Accuracies on ImageNet: Top1=74.5%, Top5=92%.

Resnet-50 (2015)

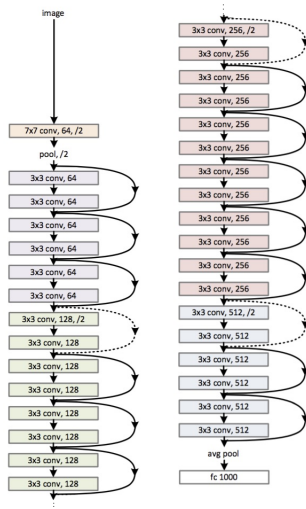




Identity function is easy to learn for residual block.

Resnet-50 (2015)

34-layer residual



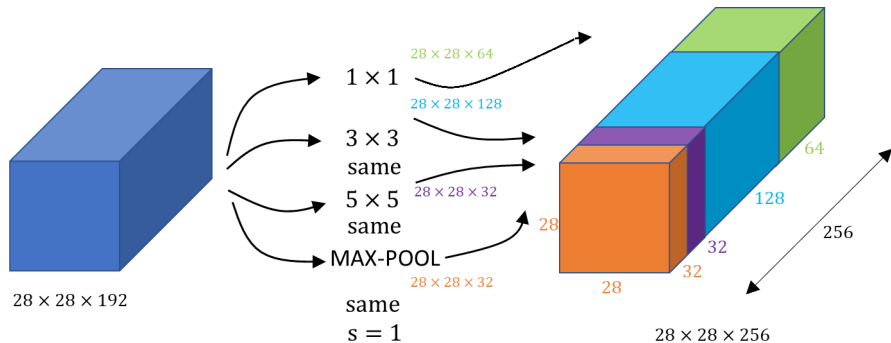
- Resnet-50 has approximately 25.6M parameters.

Resnet-50 (2015)

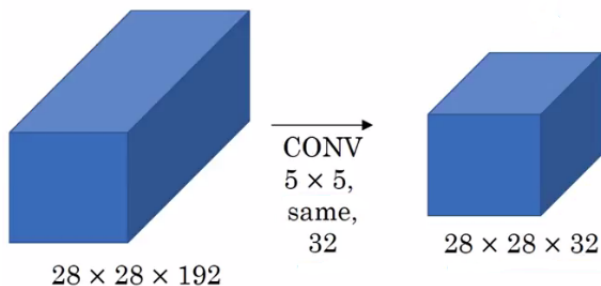
- Resnet-50 has approximately 25.6M parameters.
- Accuracies on ImageNet: Top1=77.15%, Top5=93.29%.

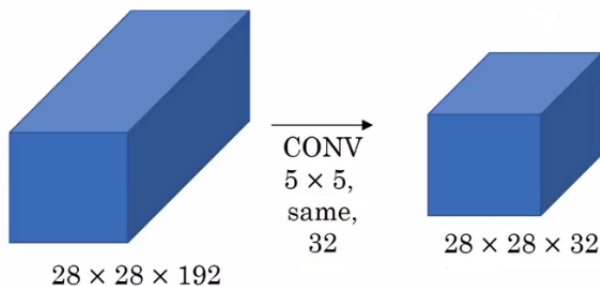
GoogLeNet/Inception v1 (2014)

GoogLeNet/Inception v1 (2014)

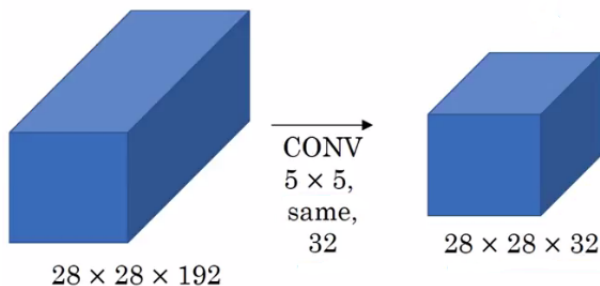


GoogLeNet/Inception v1 (2014)





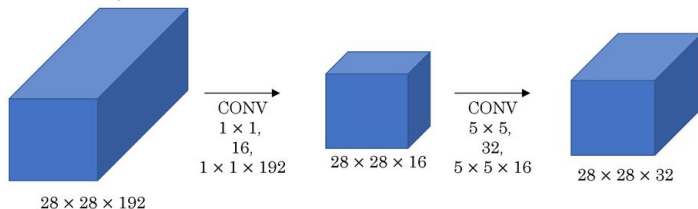
How much is the number of multiplications?



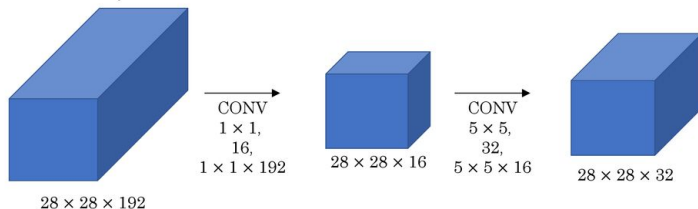
How much is the number of multiplications?

$$28 \cdot 28 \cdot 32 \cdot 5 \cdot 5 \cdot 192 \approx 120M$$

Inception Networks with 1X1 Convolutions

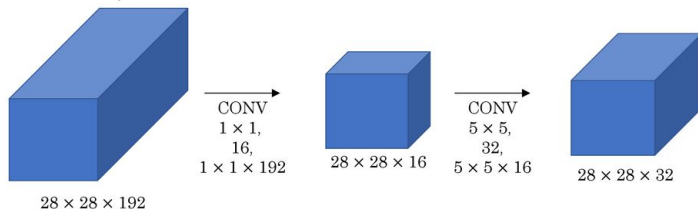


Inception Networks with 1X1 Convolutions



How much is the number of multiplications?

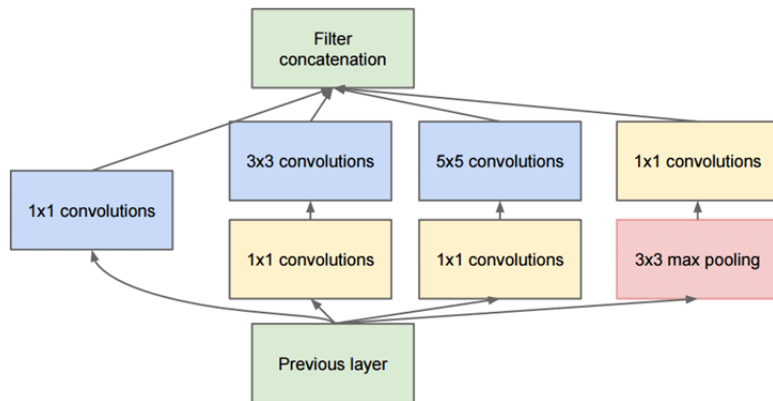
Inception Networks with 1X1 Convolutions



How much is the number of multiplications?

$$28 \cdot 28 \cdot 16 \cdot 1 \cdot 1 \cdot 192 + 28 \cdot 28 \cdot 32 \cdot 5 \cdot 5 \cdot 16 \approx 12.4M$$

GoogLeNet/Inception v1 (2014)



GoogLeNet/Inception v1 (2014)

