CNN 아키텍처

박 경 규
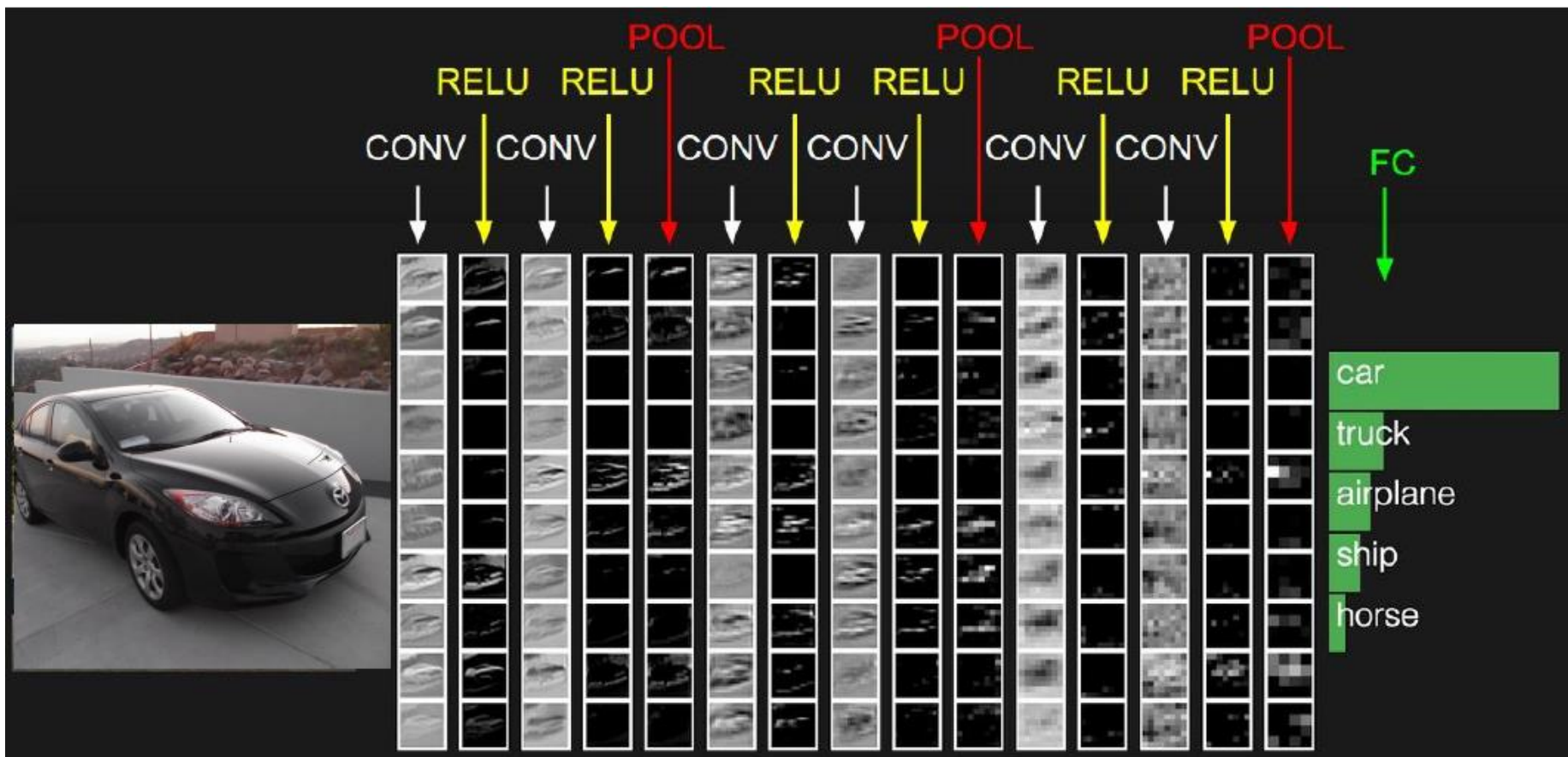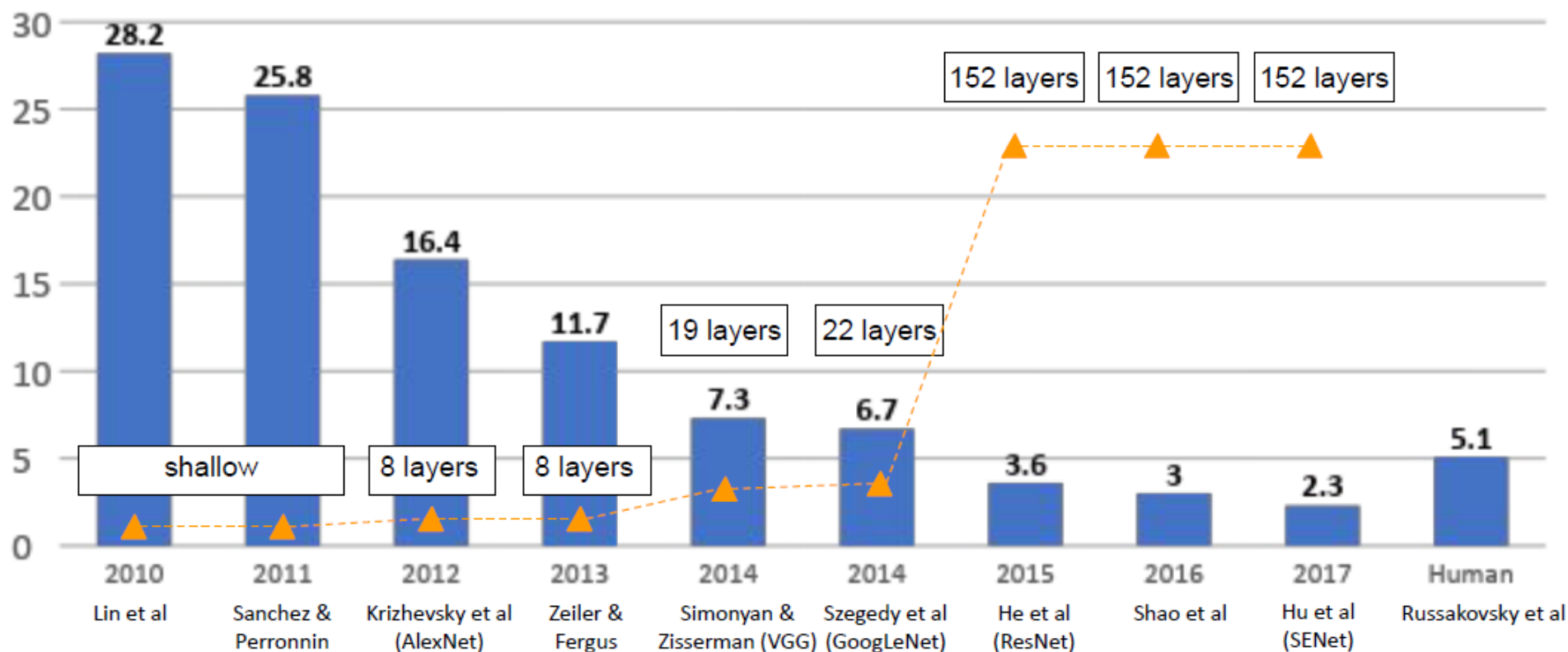
ImageNet
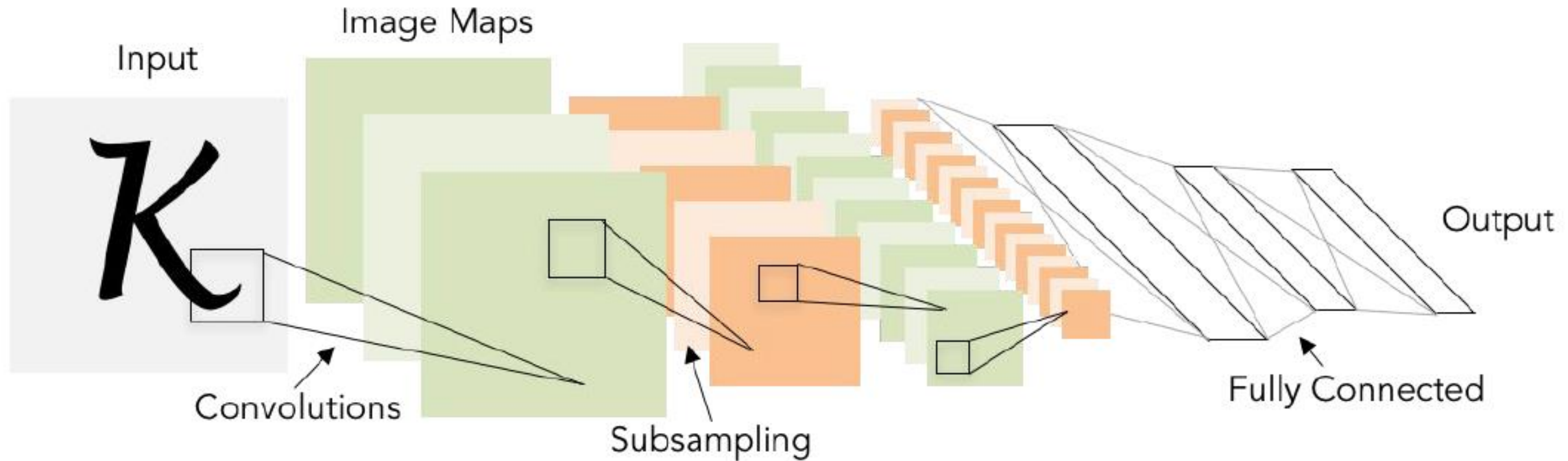
이미지 분류 모델을 측정하기 위한 데이터셋, 학습데이터셋 138G, 2만개 이상의 클래스, 약 1,400만장의 이미지

ImageNet Large Scale Visual Recognition Challenge (ILSVRC) winners
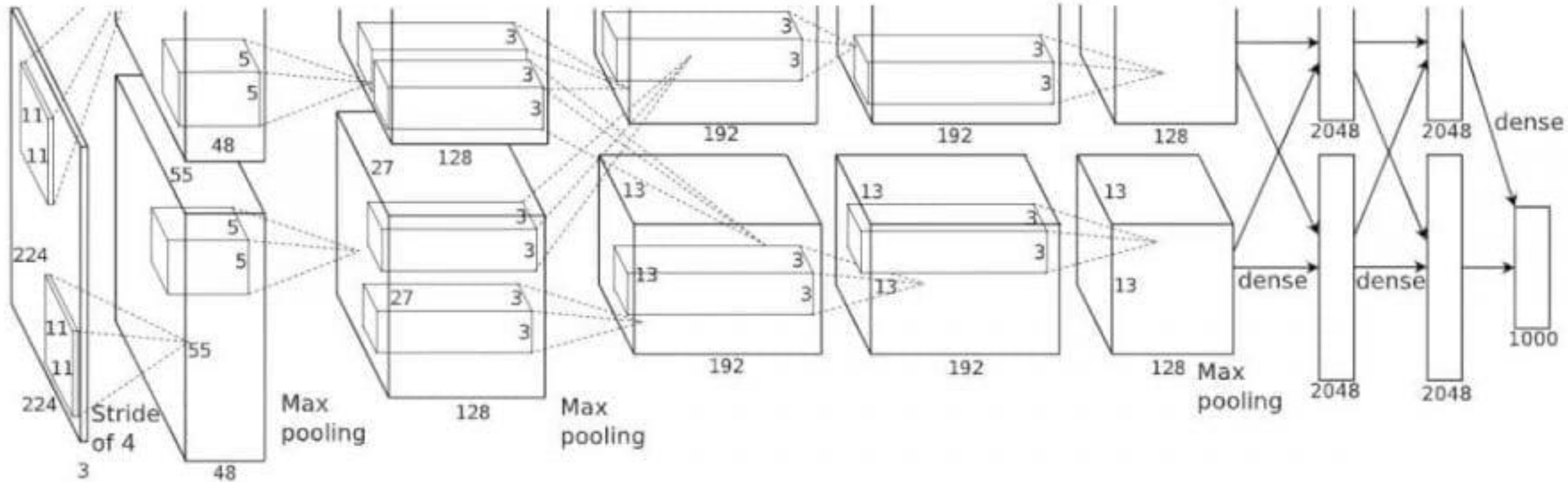
LeNet-5

Image Maps

Input
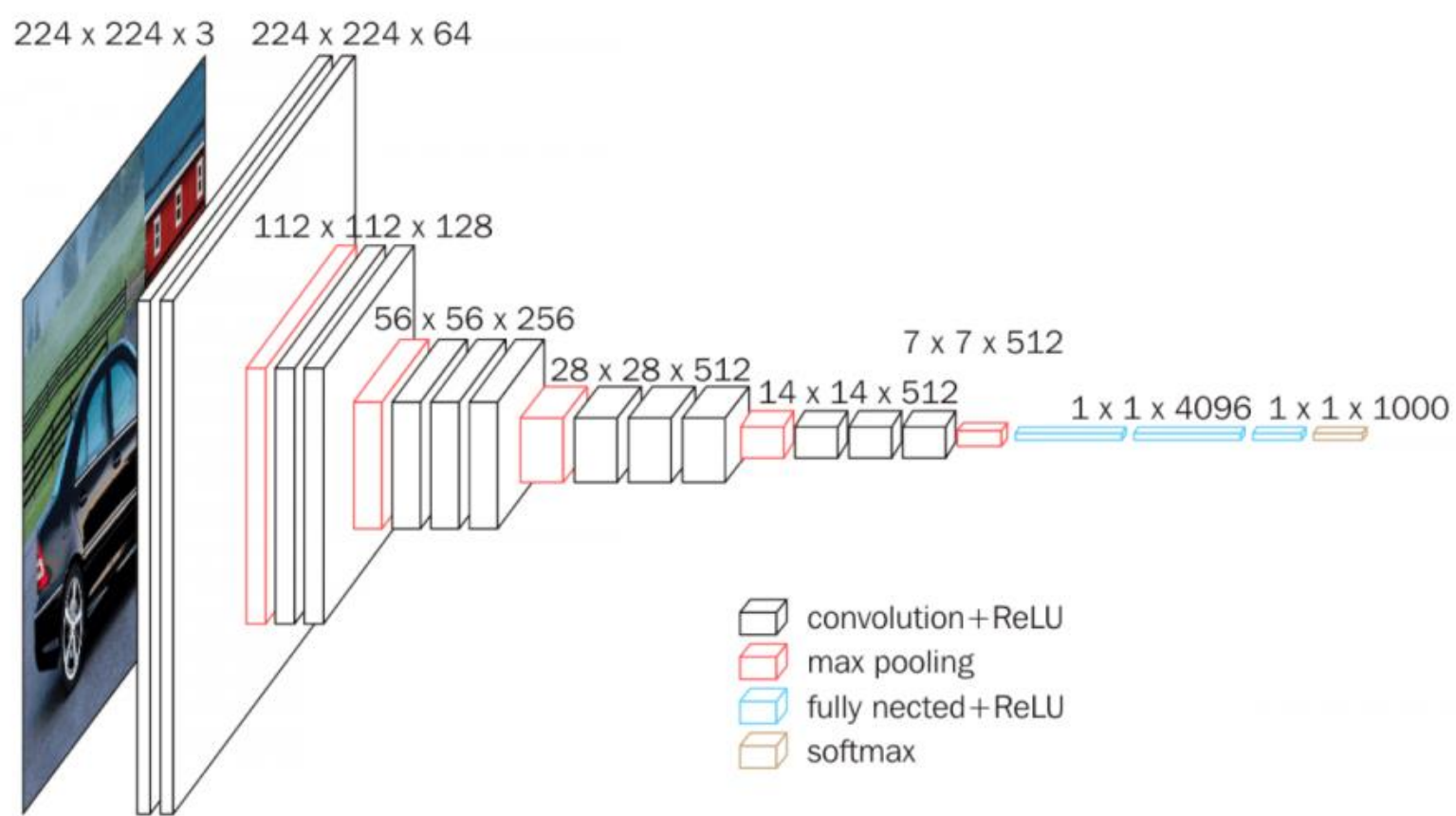

Output

Convolutions

Subsampling

Fully Connected

Conv filters were 5x5, applied at stride 1
Subsampling (Pooling) layers were 2x2 applied at stride 2
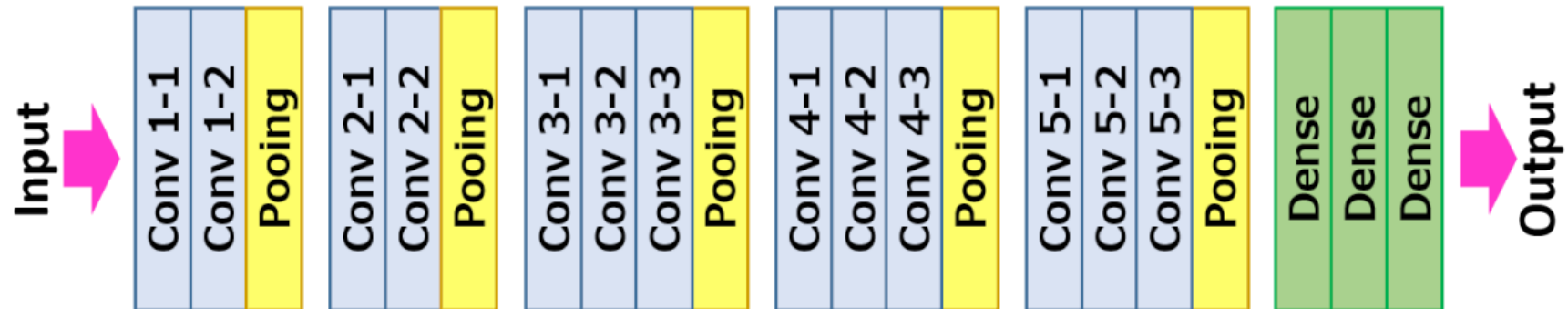i.e. architecture is [CONV-POOL-CONV-POOL-FC-FC]

AlexNet

VGG

224 x 224 x 3   224 x 224 x 64

112 x 112 x 128

56 x 56 x 256

28 x 28 x 512

14 x 14 x 512

7 x 7 x 512

1 x 1 x 4096   1 x 1 x 1000

- convolution+ReLU
- max pooling
- fully nected+ReLU
- softmax

VGG-16

Input → Conv 1-1 | Conv 1-2 | Pooing → Conv 2-1 | Conv 2-2 | Pooing → Conv 3-1 | Conv 3-2 | Conv 3-3 | Pooing → Conv 4-1 | Conv 4-2 | Conv 4-3 | Pooing → Conv 5-1 | Conv 5-2 | Conv 5-3 | Pooing → Dense | Dense | Dense → Output

https://neurohive.io/en/popular-networks/vgg16/

5

VGG

ConvNet Configuration					
A	**A-LRN**	**B**	**C**	**D**	**E**
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

GoogLeNet

[Szegedy et al., 2014]

Full GoogLeNet
architecture
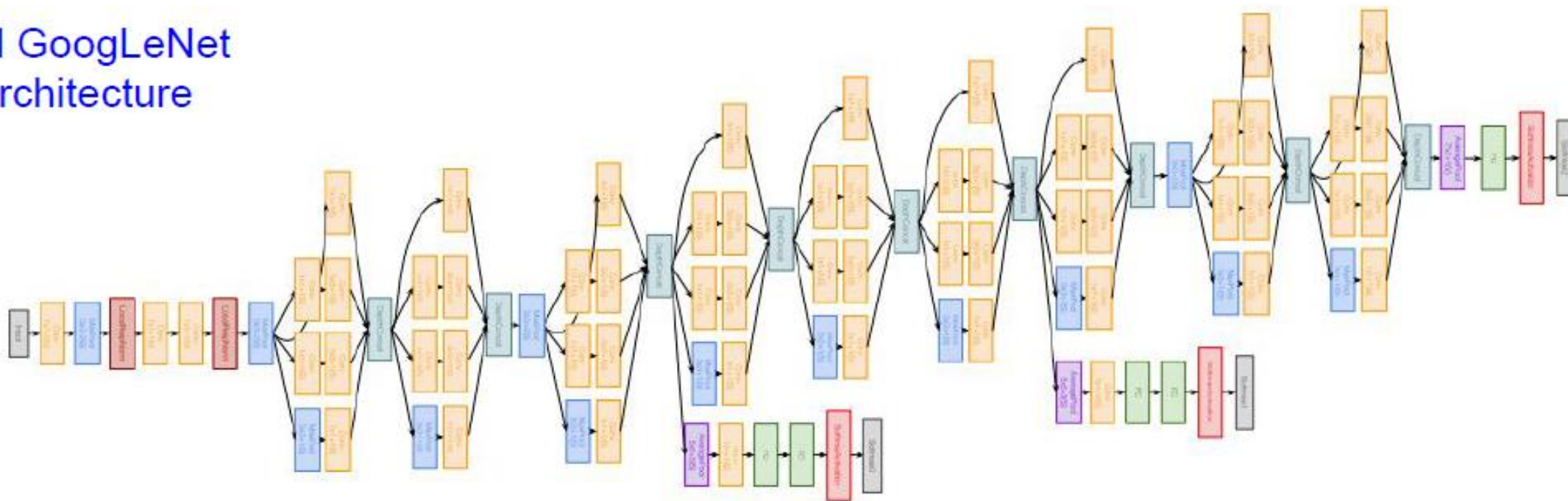
22 total layers with weights
(parallel layers count as 1 layer => 2 layers per Inception module. Don't count auxiliary output layers)

GoogLeNet

Deeper networks, with computational efficiency

- 22 layers
- Efficient "Inception" module
- Avoids expensive FC layers
- 12x less params than AlexNet
- 27x less params than VGG-16
- ILSVRC'14 classification winner (6.7% top 5 error)
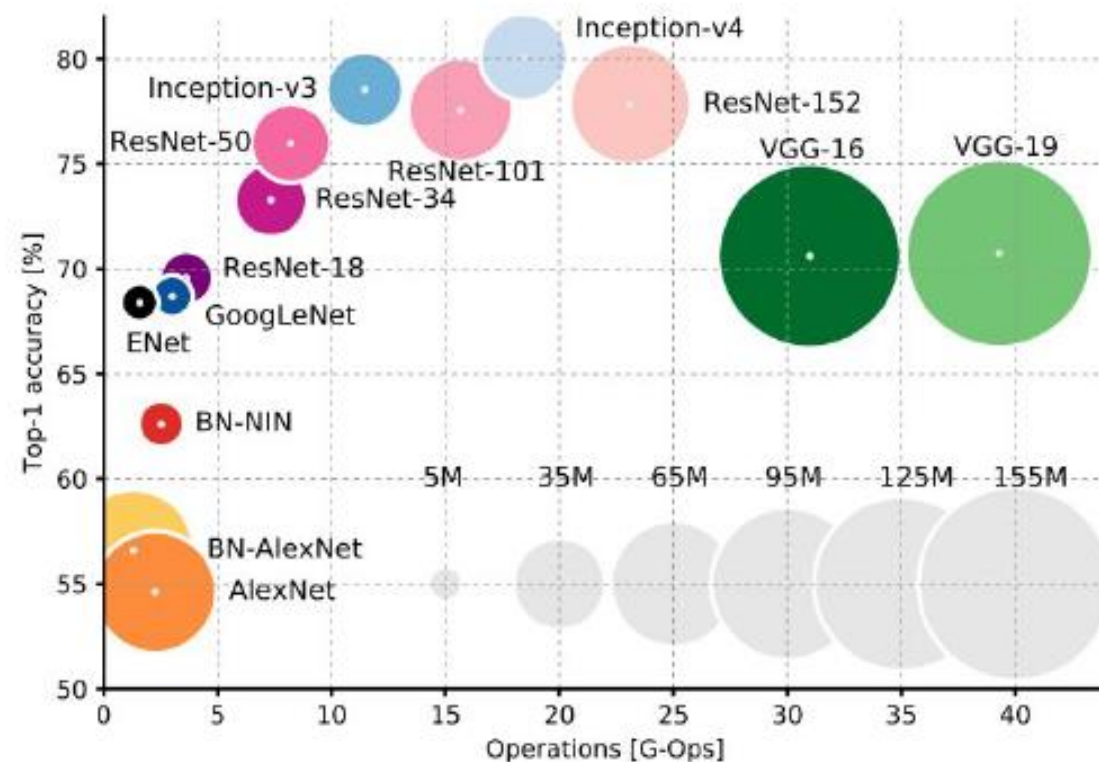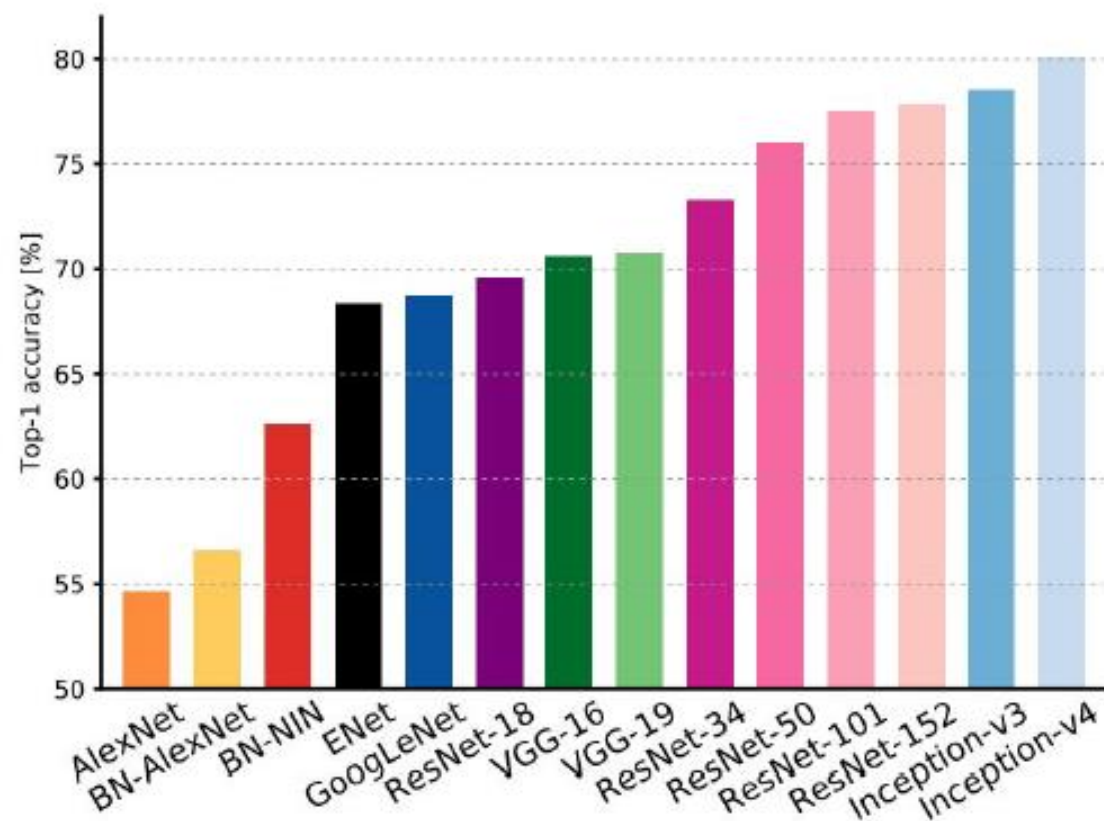
Inception module
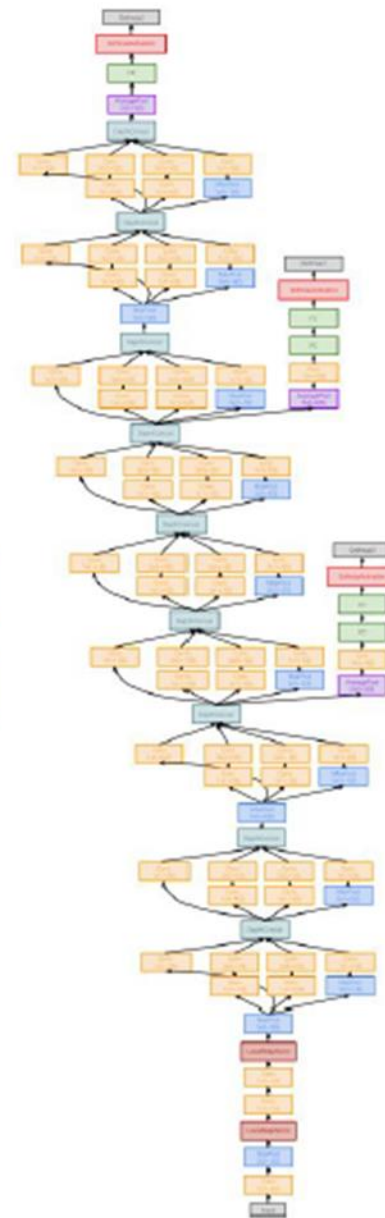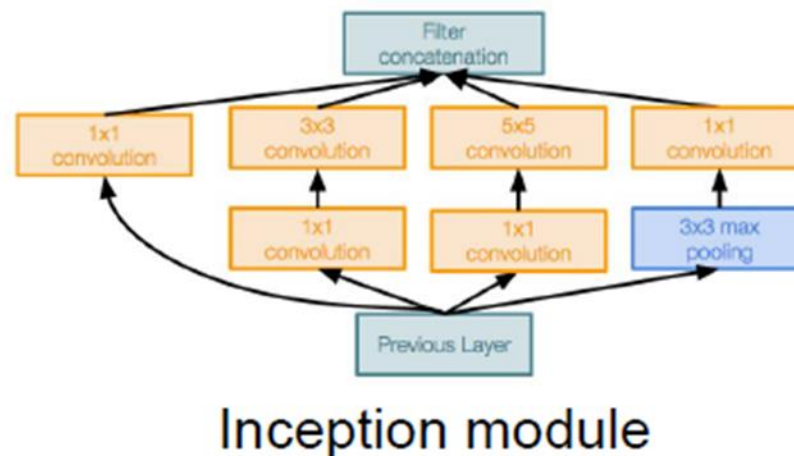
ResNet

Very deep networks using residual connections

- 152-layer model for ImageNet
- ILSVRC'15 classification winner (3.57% top 5 error)
- Swept all classification and detection competitions in ILSVRC'15 and COCO'15!

$F(x) + x$

relu

conv

$F(x)$ relu

conv

X identity

X

Residual block
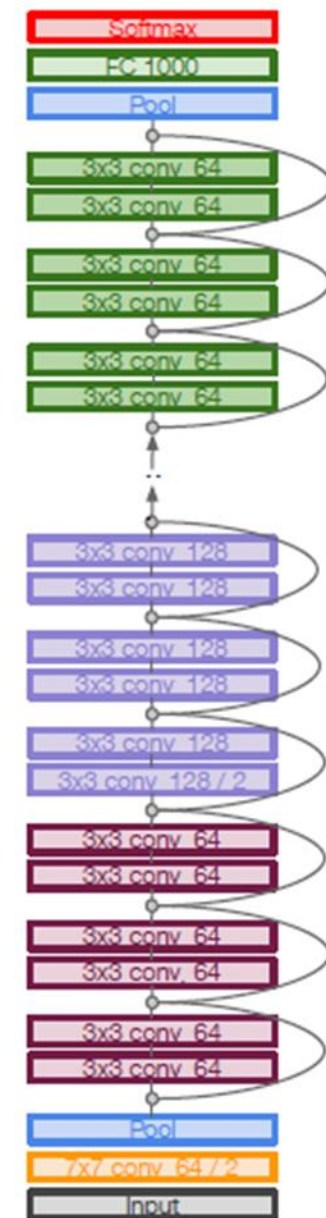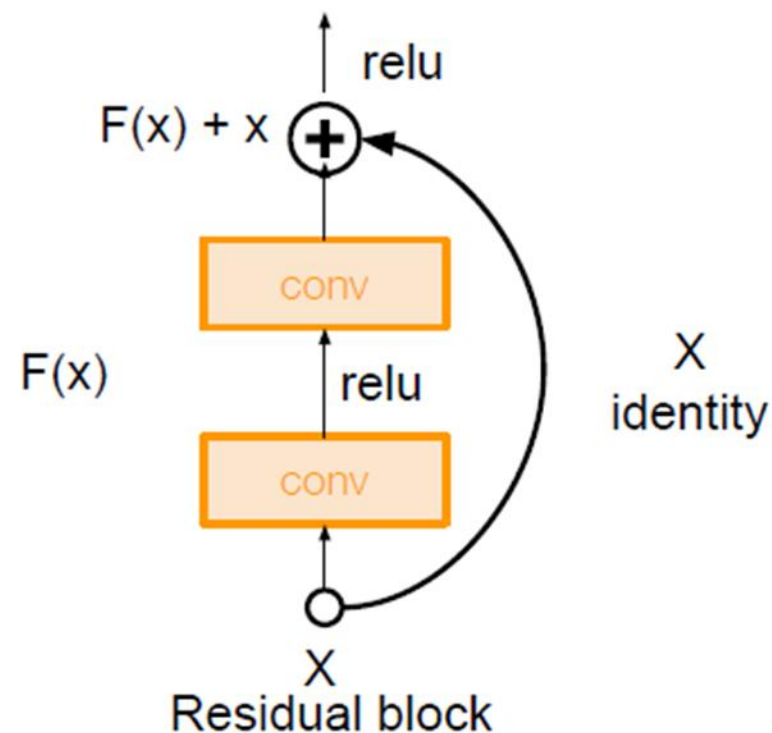
Comparing complexity...

An Analysis of Deep Neural Network Models for Practical Applications, 2017.

Further Study

CS231n

Lecture 9: CNN Architecture

Thank you

kgpark88@gmail.com