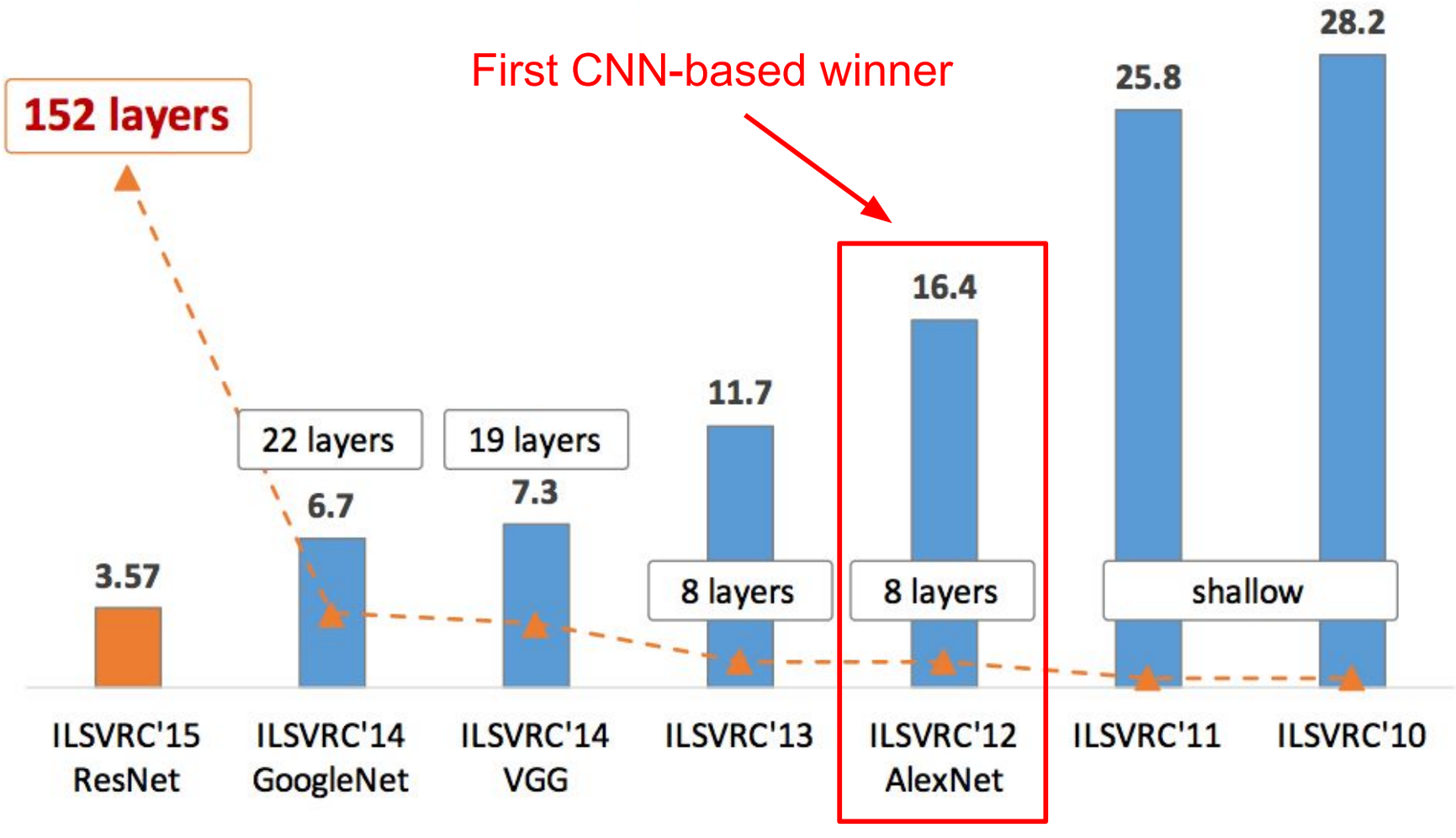
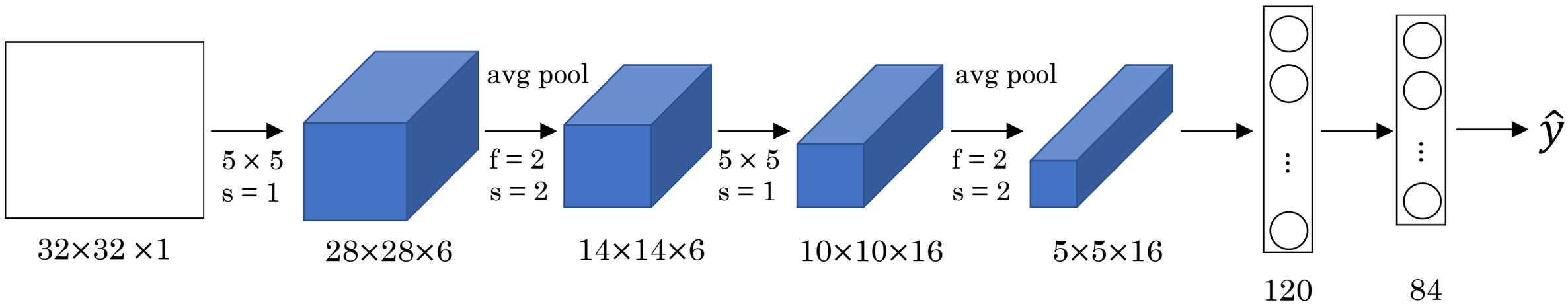
# CNN Architectures

# ImageNet Large Scale Visual Recognition Challenge (ILSVRC) winners

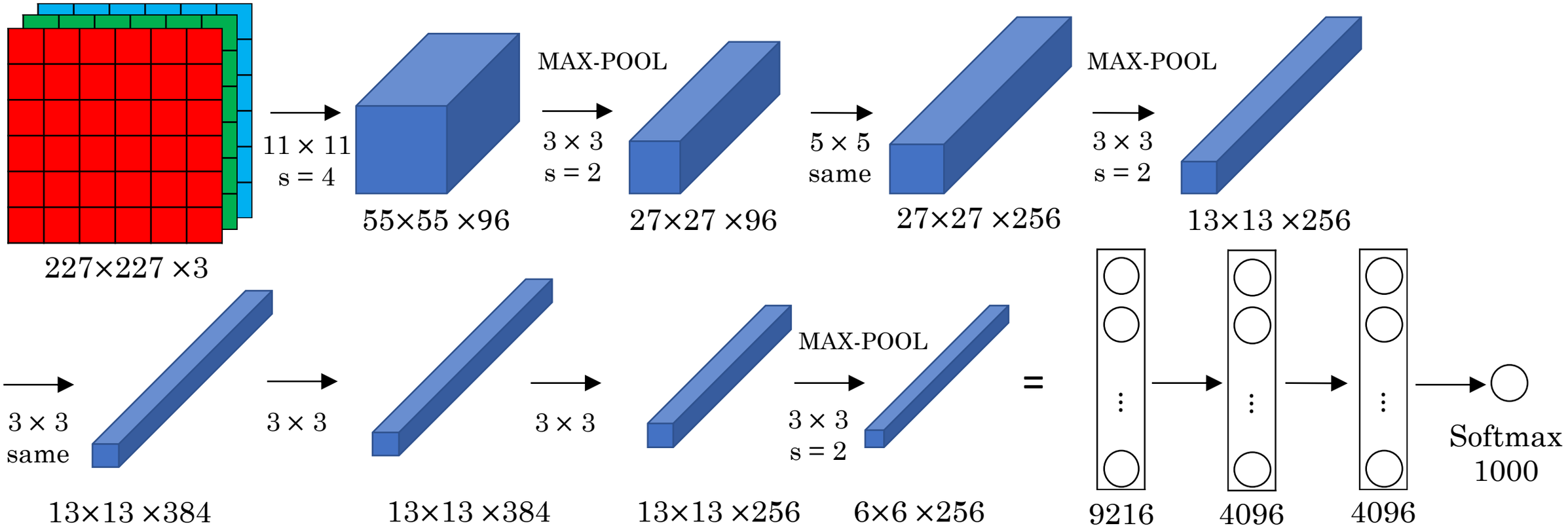


# LeNet




[LeCun et al., 1998. Gradient-based learning applied to document recognition]

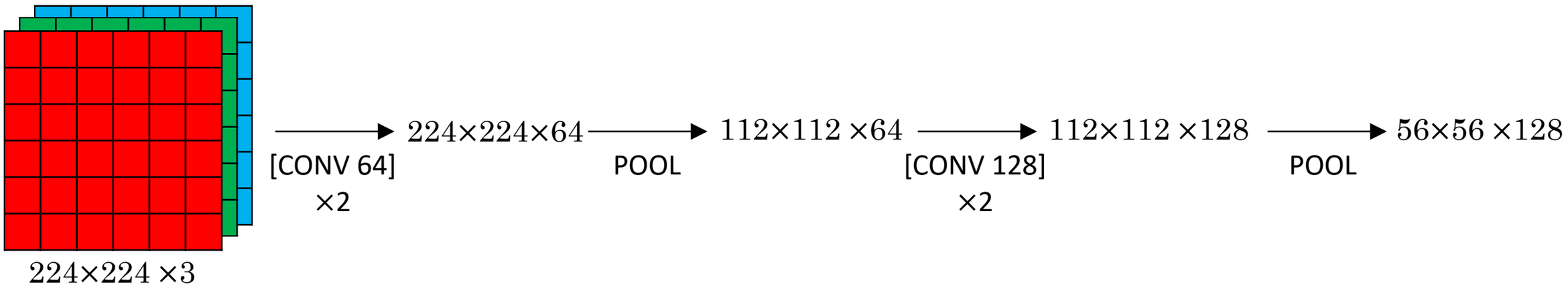
# AlexNet




[Krizhevsky et al., 2012. ImageNet classification with deep convolutional neural networks]

# VGG

CONV = 3×3 filter, s = 1, same  MAX-POOL = 2×2 , s = 2




[Simonyan & Zisserman 2015. Very deep convolutional networks for large-scale image recognition]

# Inception Net


http://knowyourmeme.com/memes/we-need-to-go-deeper

# Why does a 1 × 1 convolution do?

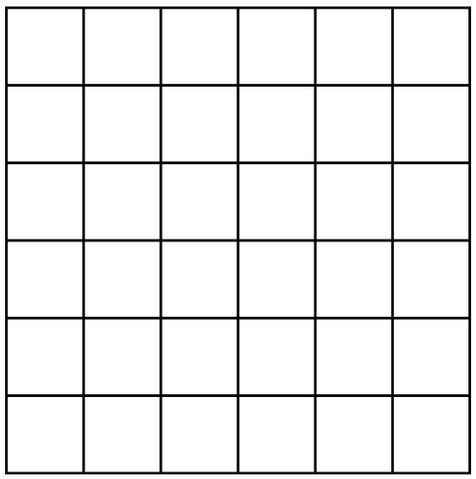
| | | | | | |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 6 | 5 | 8 |
| 3 | 5 | 5 | 1 | 3 | 4 |
| 2 | 1 | 3 | 4 | 9 | 3 |
| 4 | 7 | 8 | 5 | 7 | 9 |
| 1 | 5 | 3 | 7 | 4 | 8 |
| 5 | 4 | 9 | 8 | 3 | 5 |

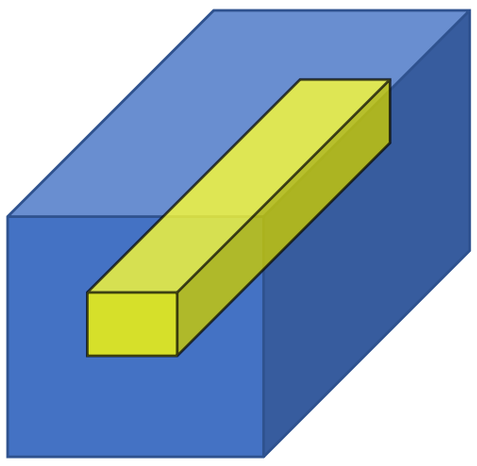
6 × 6

*

2

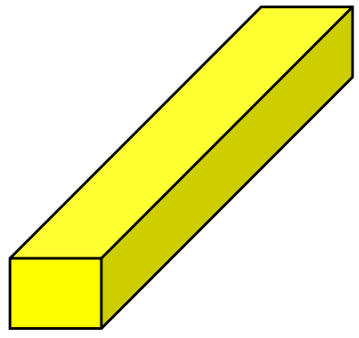
=





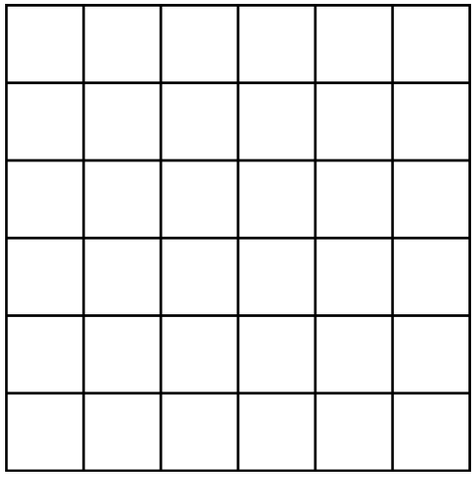
6 × 6 × 32

*



1 × 1 × 32

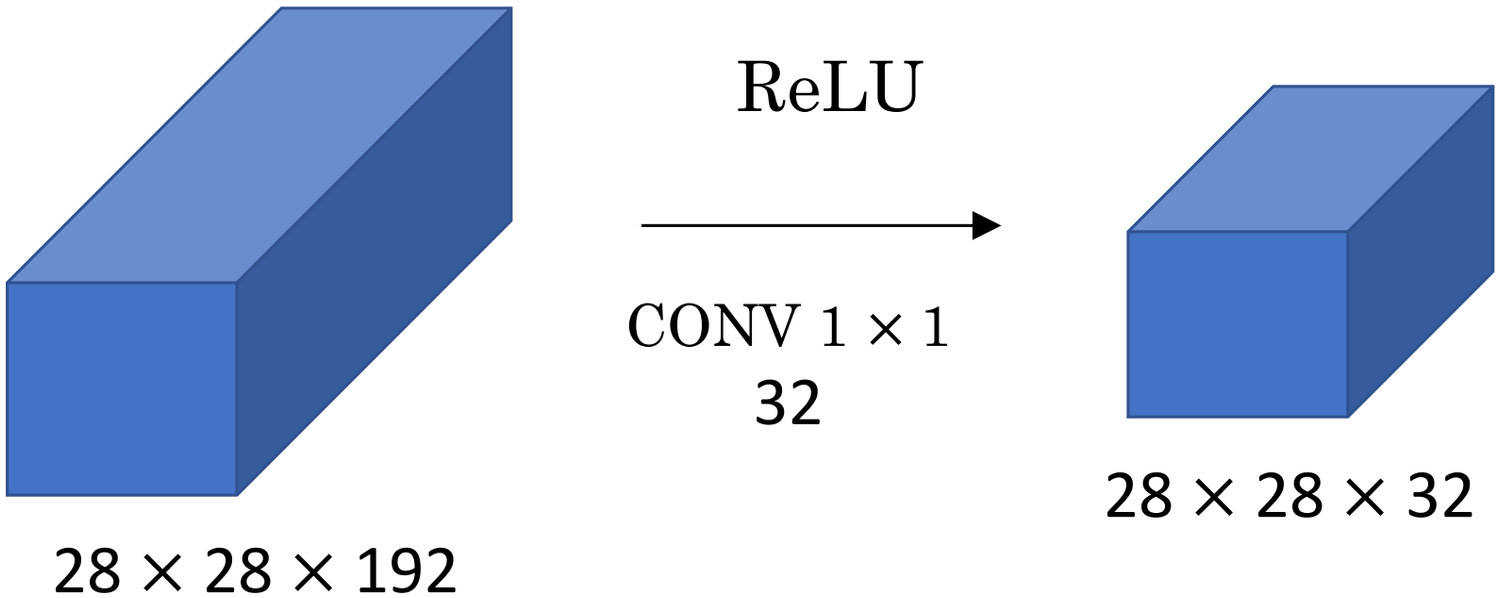
=



6 × 6 × # filters

[Lin et al., 2013. Network in network]

# Using 1×1 convolutions

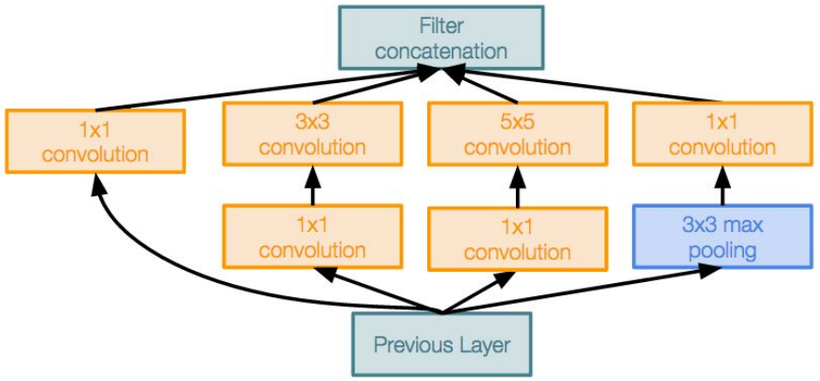


# Inception Module

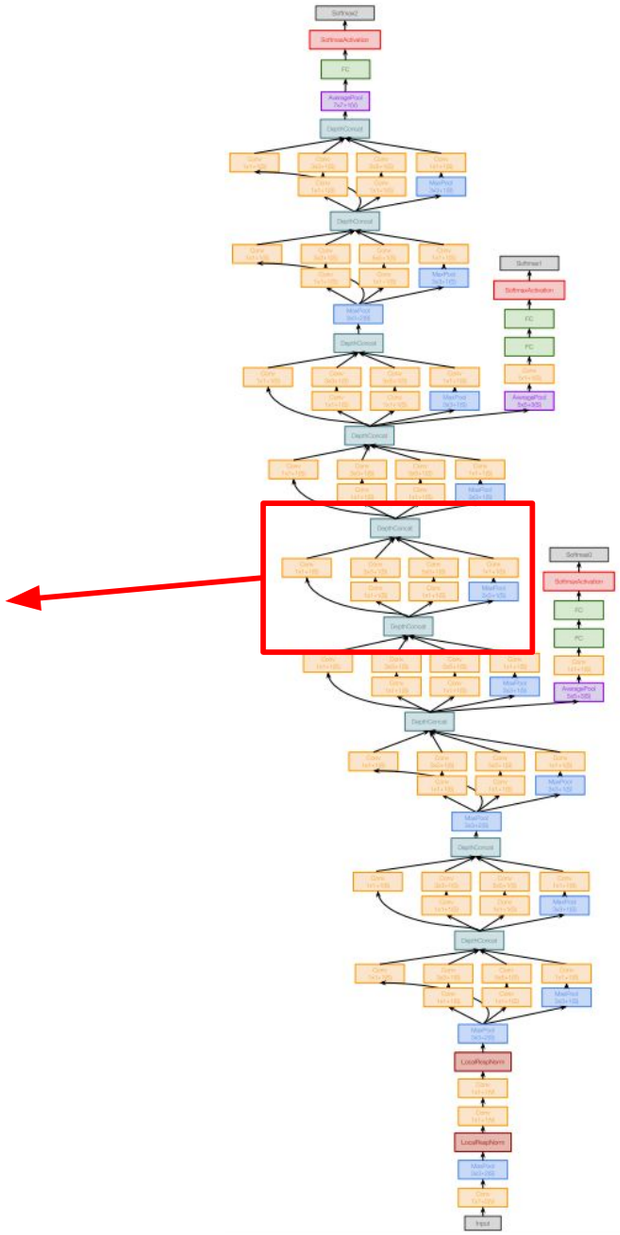
## Case Study: GoogLeNet

*[Szegedy et al., 2014]*

"Inception module": design a good local network topology (network within a network) and then stack these modules on top of each other

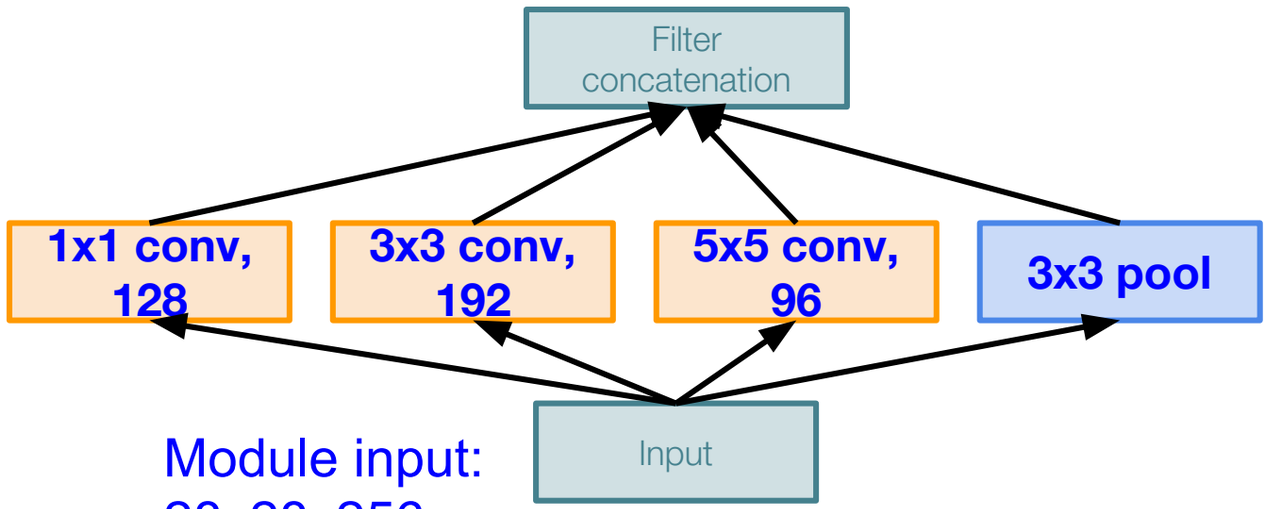



Inception module



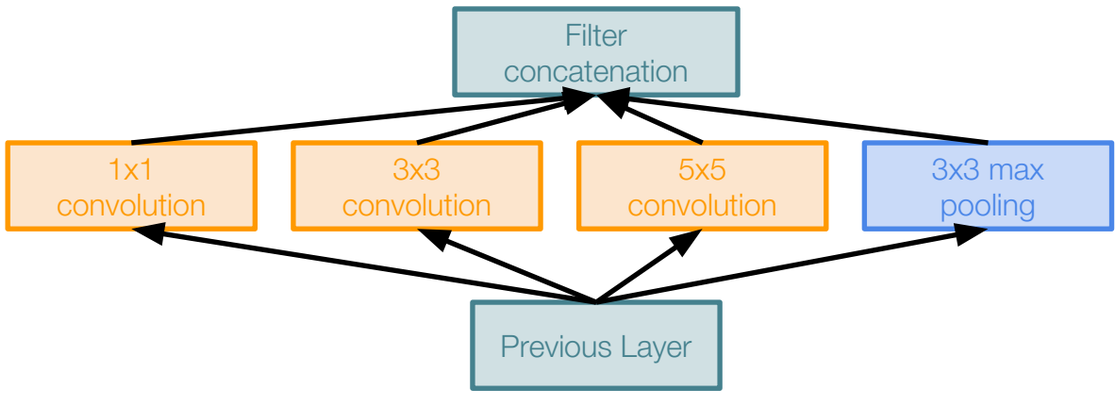
# Inception Module

Example:



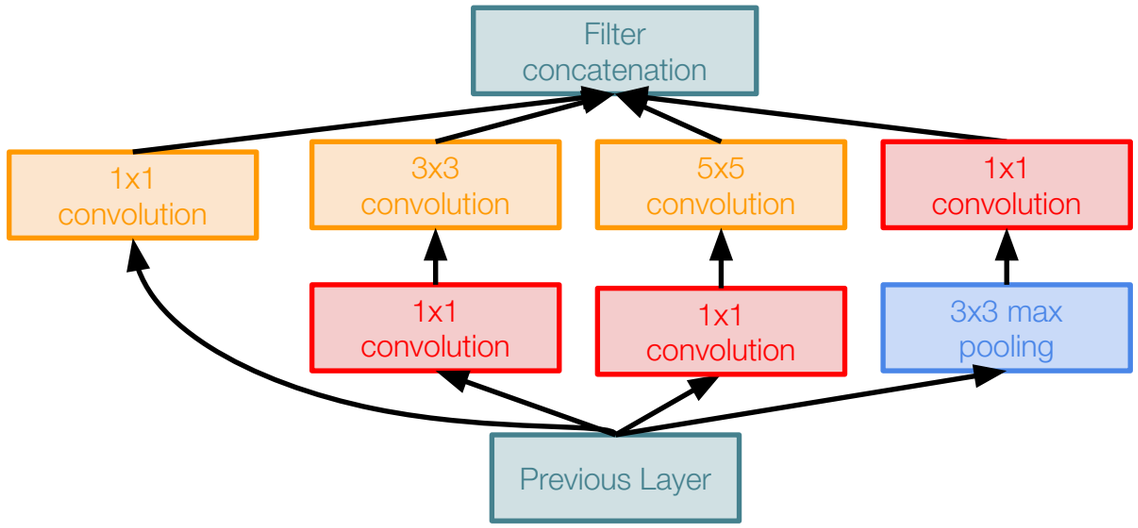

Naive Inception module

# With Bottleneck




Naive Inception module

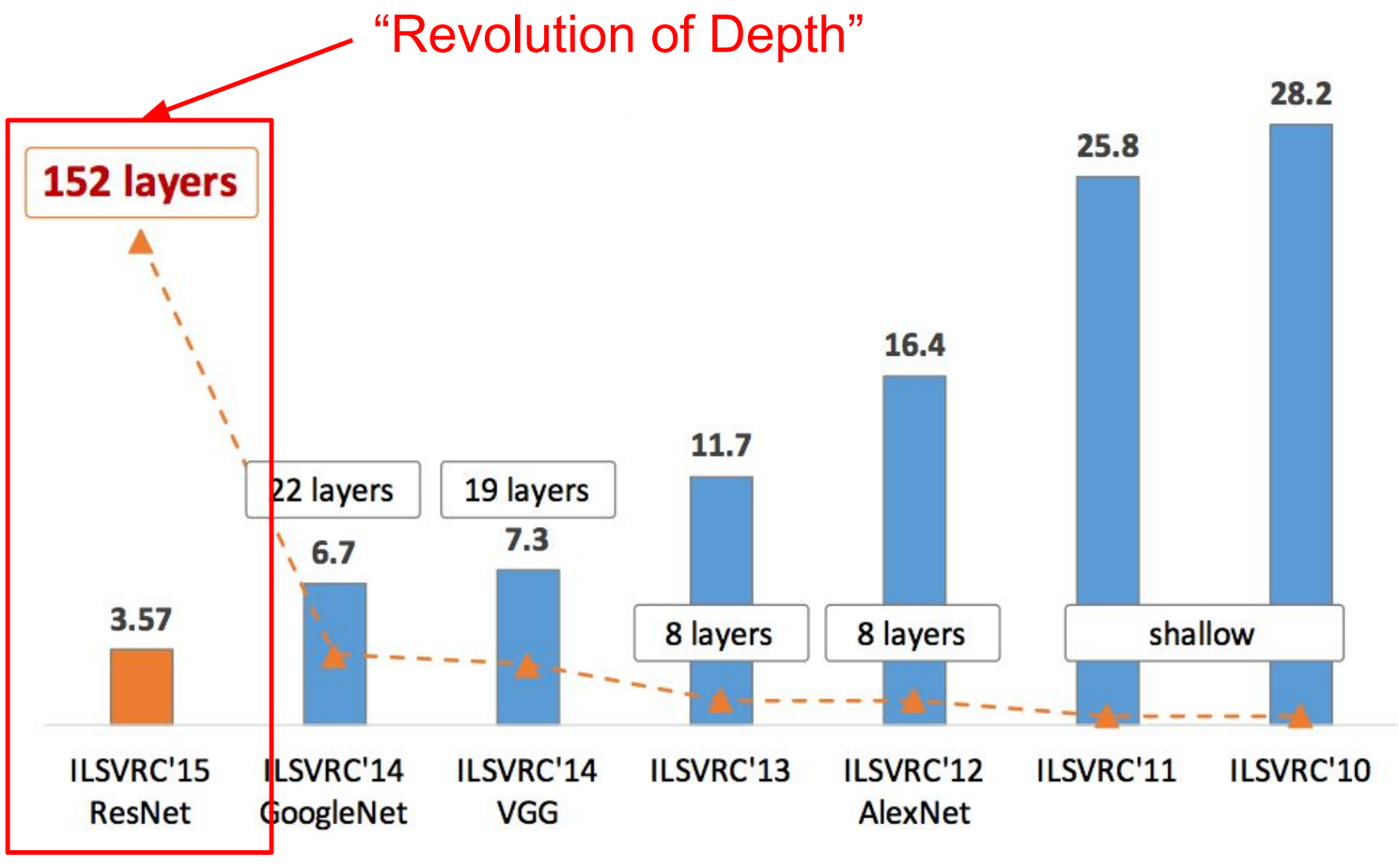
1x1 conv “bottleneck”
layers



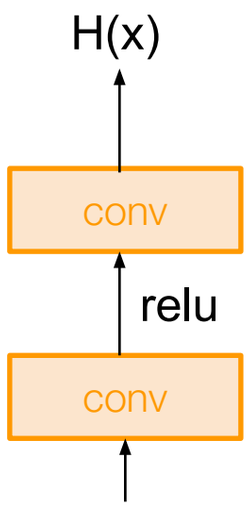

Inception module with dimension reduction

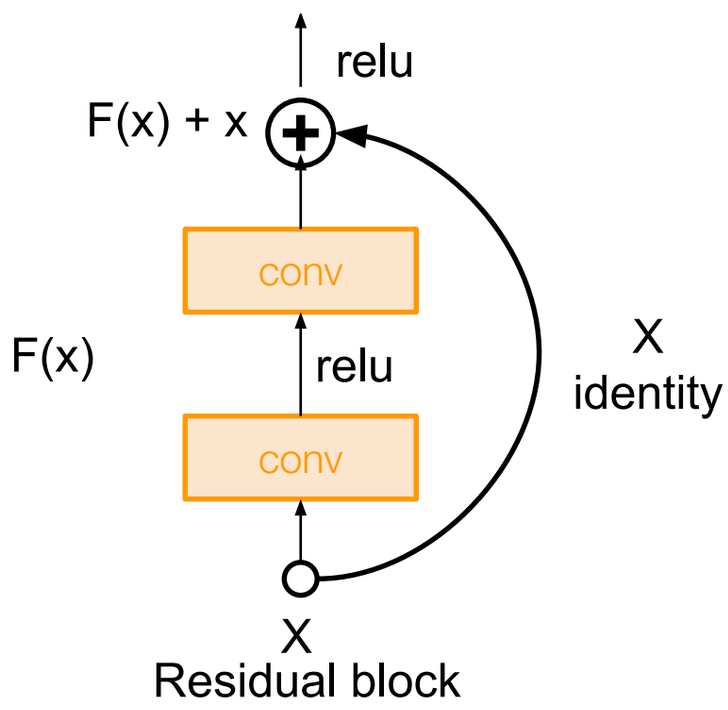
# More Deeper

ImageNet Large Scale Visual Recognition Challenge (ILSVRC) winners

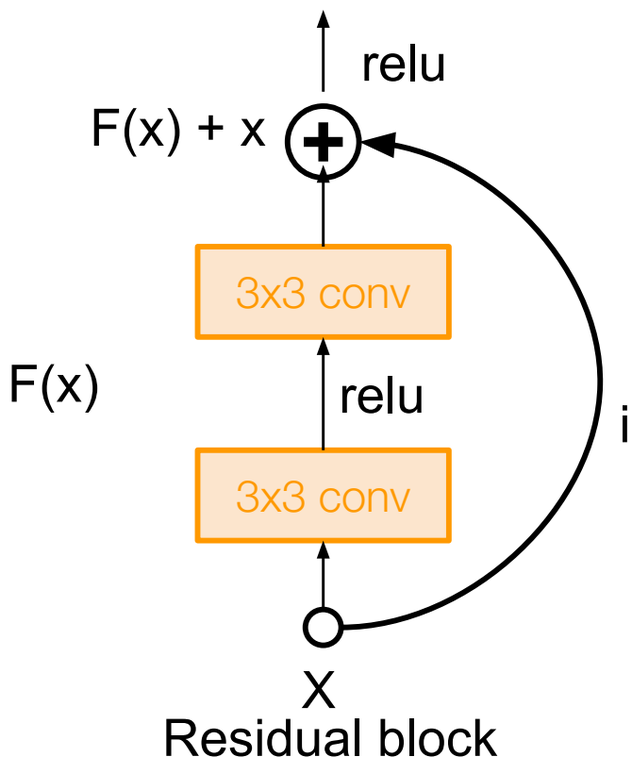


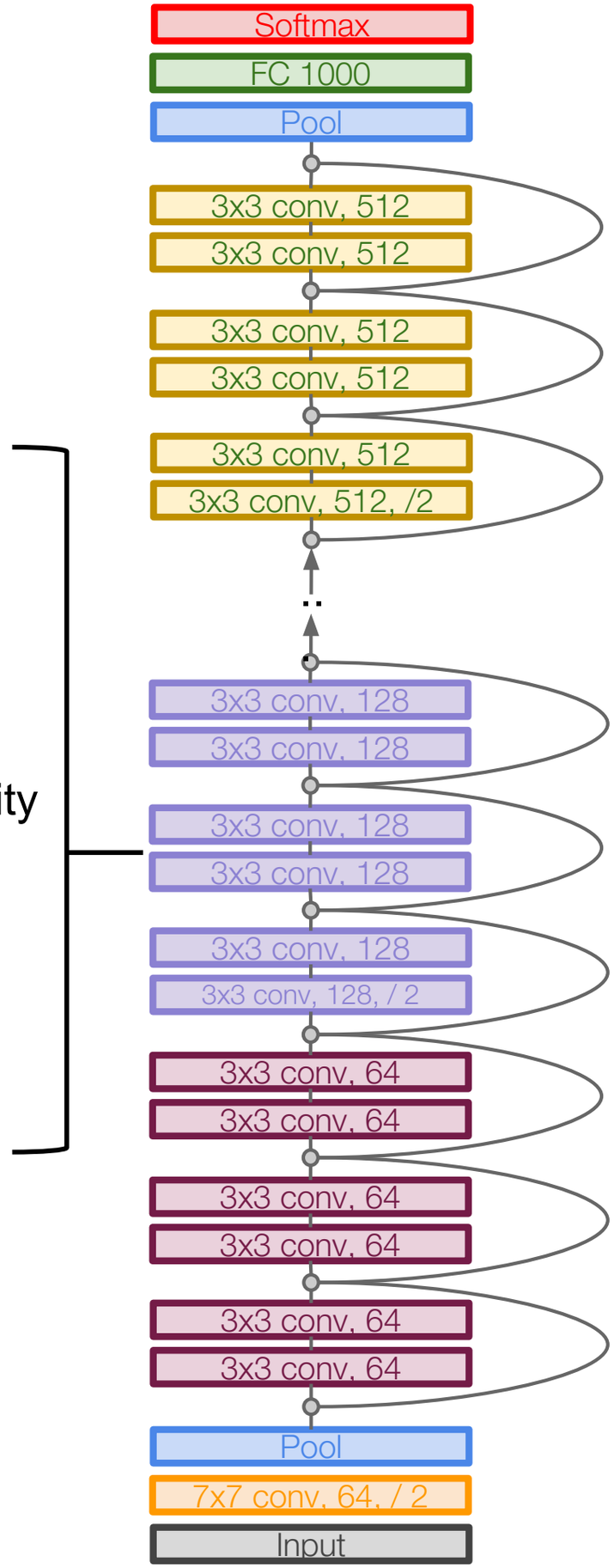
# Residual Block



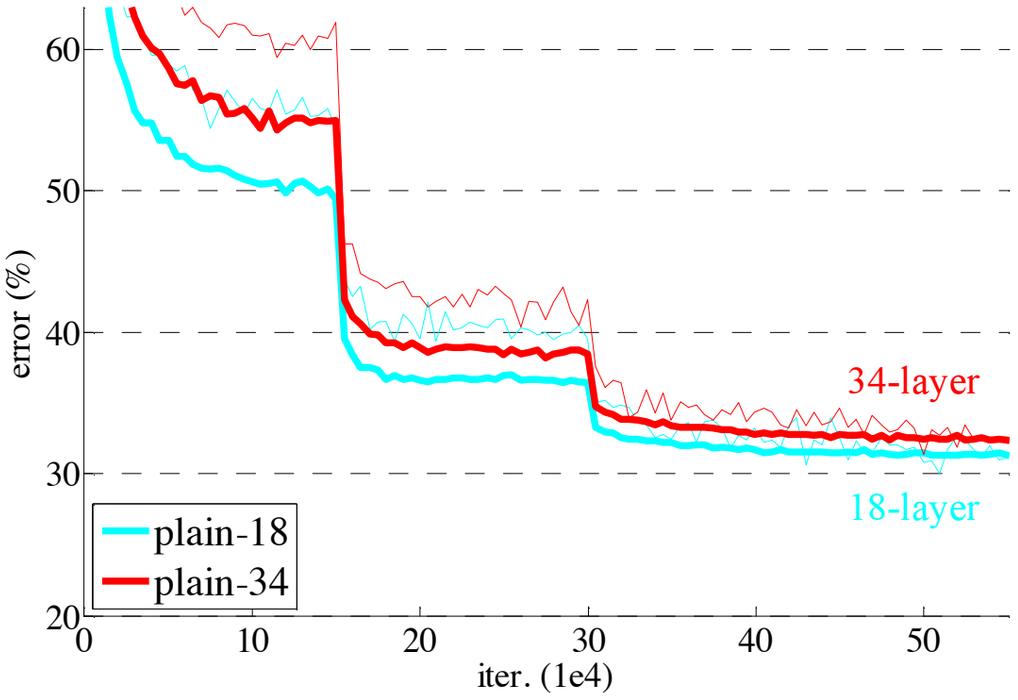



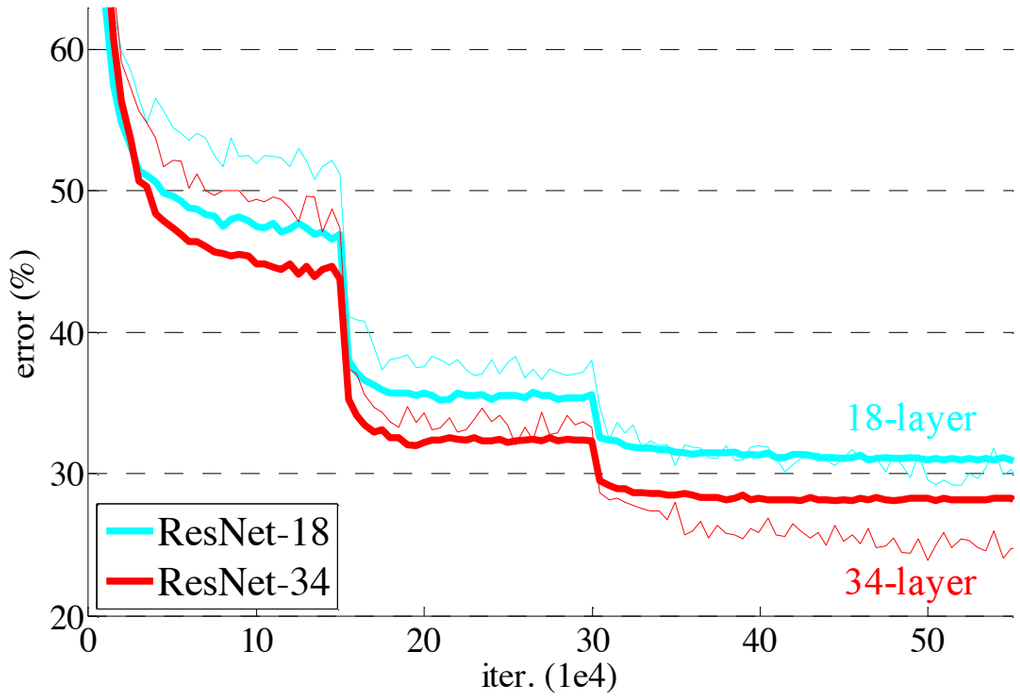
# ResNet


relu
F(x) + x
3x3 conv
F(x)
relu
3x3 conv
X
identity
X
Residual block


Softmax
FC 1000
Pool
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512, /2
..
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128, / 2
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
Pool
7x7 conv, 64, / 2
Input

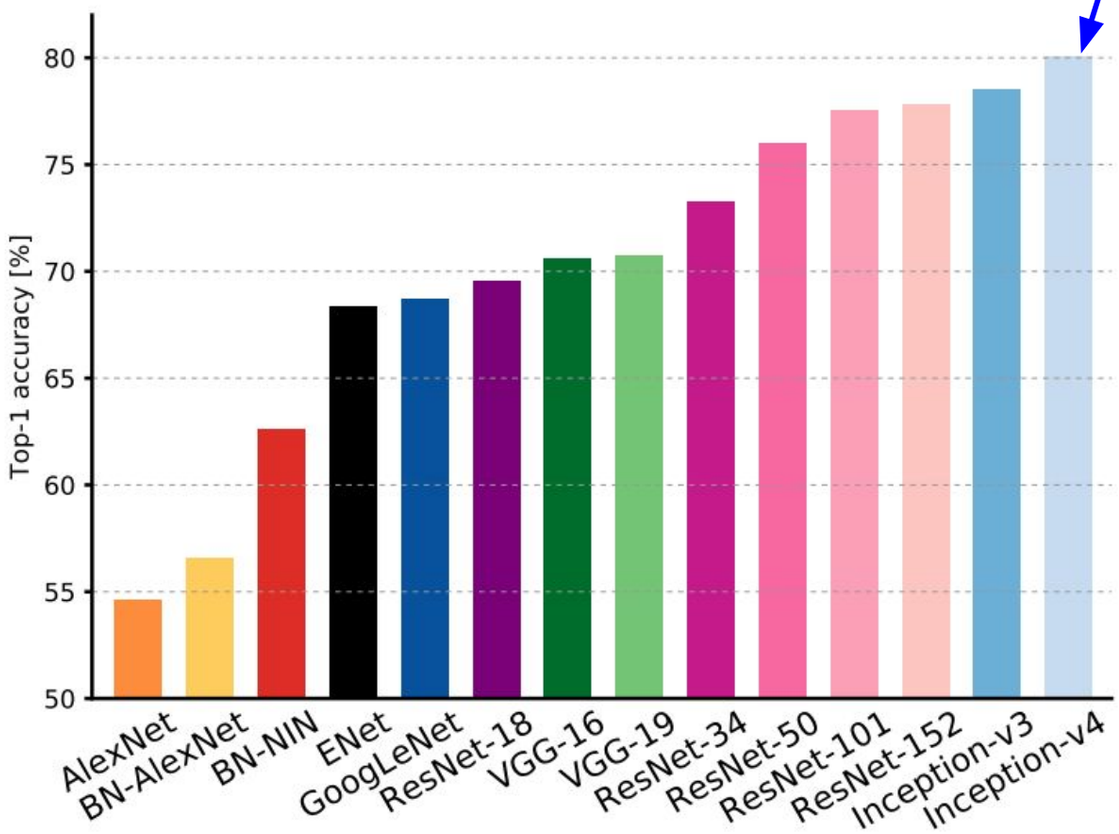
# ResNet results

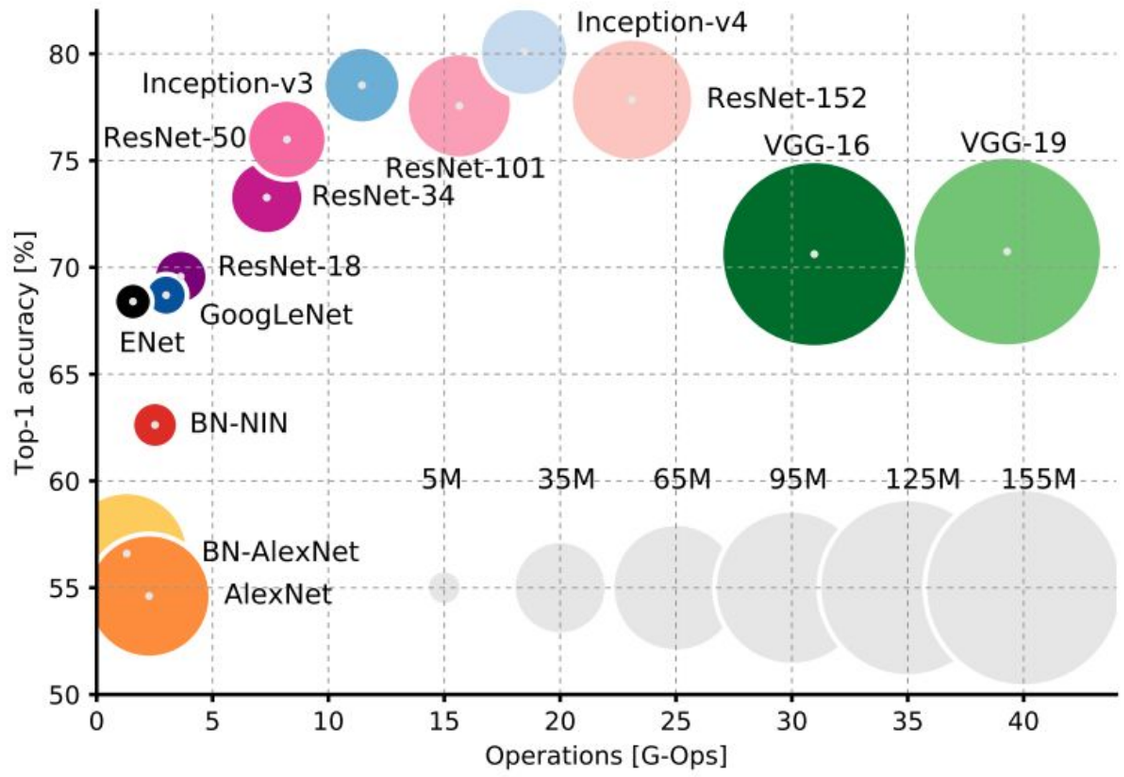


# Comparing complexity...

Inception-v4: Resnet + Inception!




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