

ResNet

Deep Residual Learning for Image Recognition

CONTENTS

01 Basic CNN

LeNet

AlexNet

VGG

GoogleNet

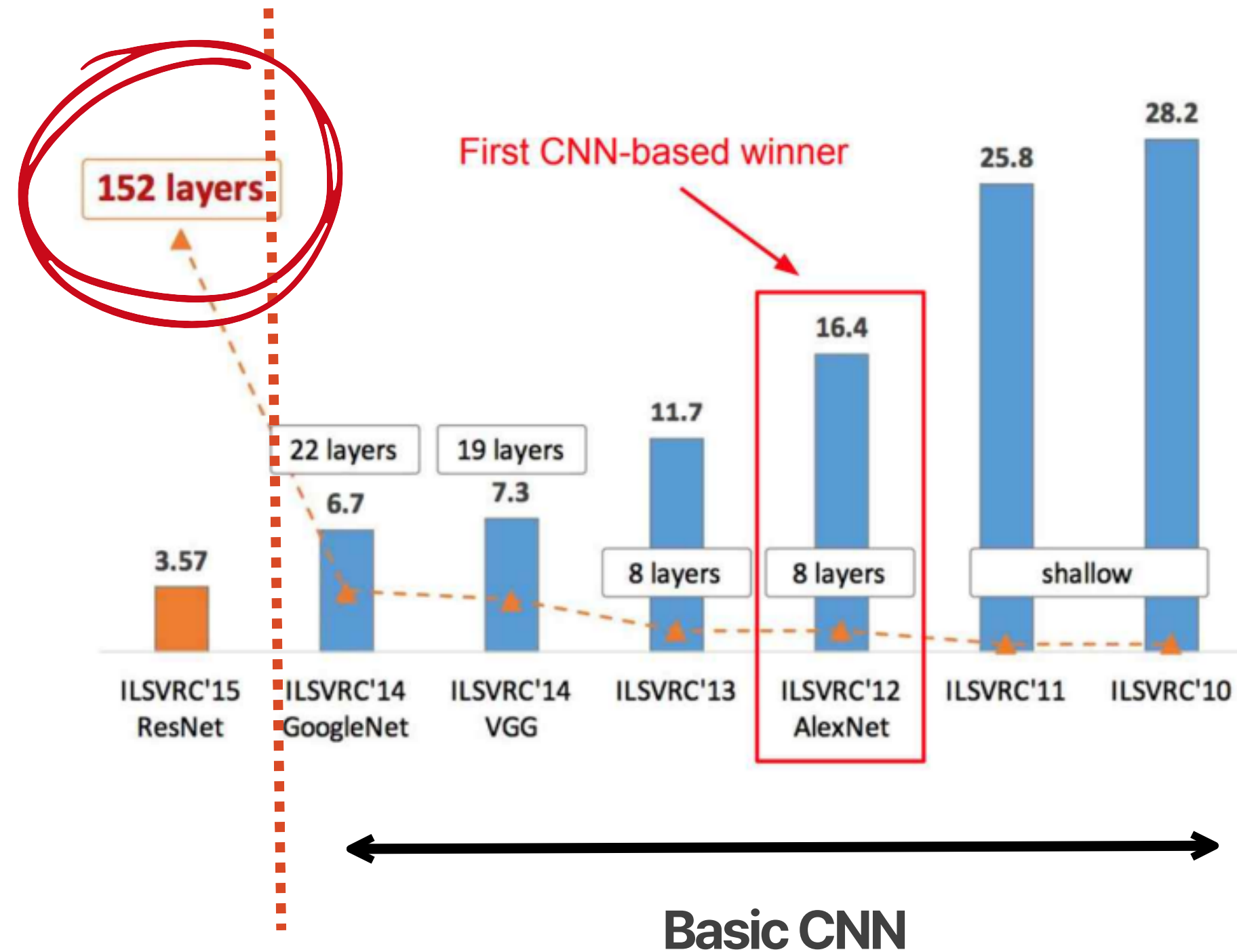
02 ResNet

Residual Learning

ImageNet 결과 분석

Bottle Neck

ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

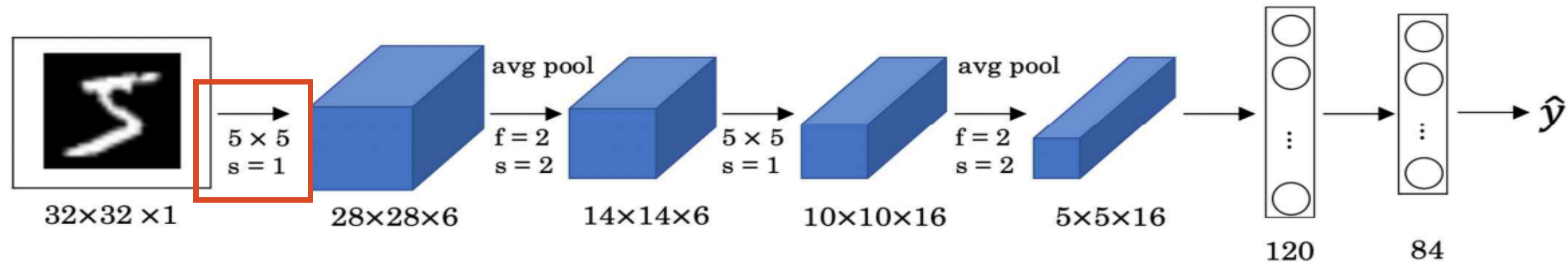


Basic CNN

LeNet

Convolution Neural Network 개념이 처음 도입된 최초의 CNN 모델

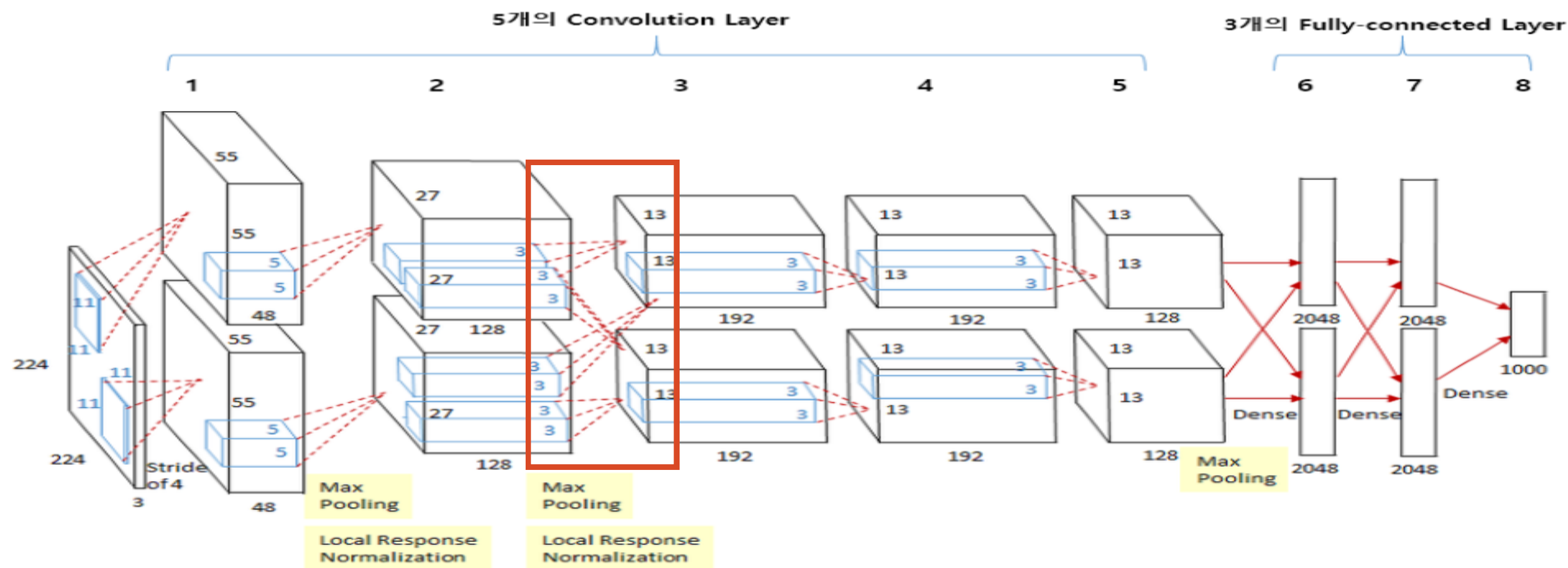
Covolution & Subsampling 반복, fully - connected



AlexNet

LeNet과 병렬적인 구조 차이

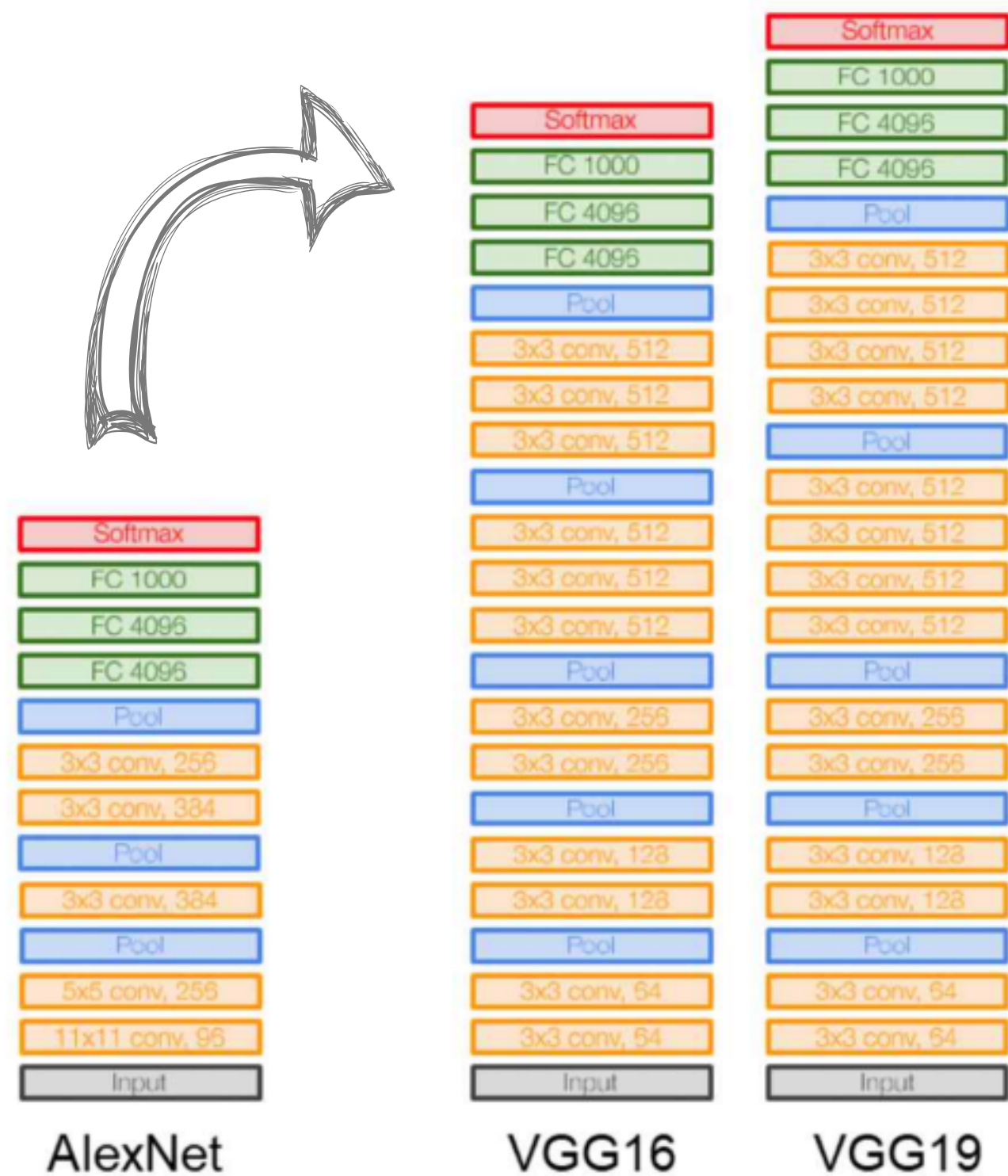
CV 딥러닝 기반 모델의 급격한 발전 시작





모든 Convolution Layer의 Kernel 3*3

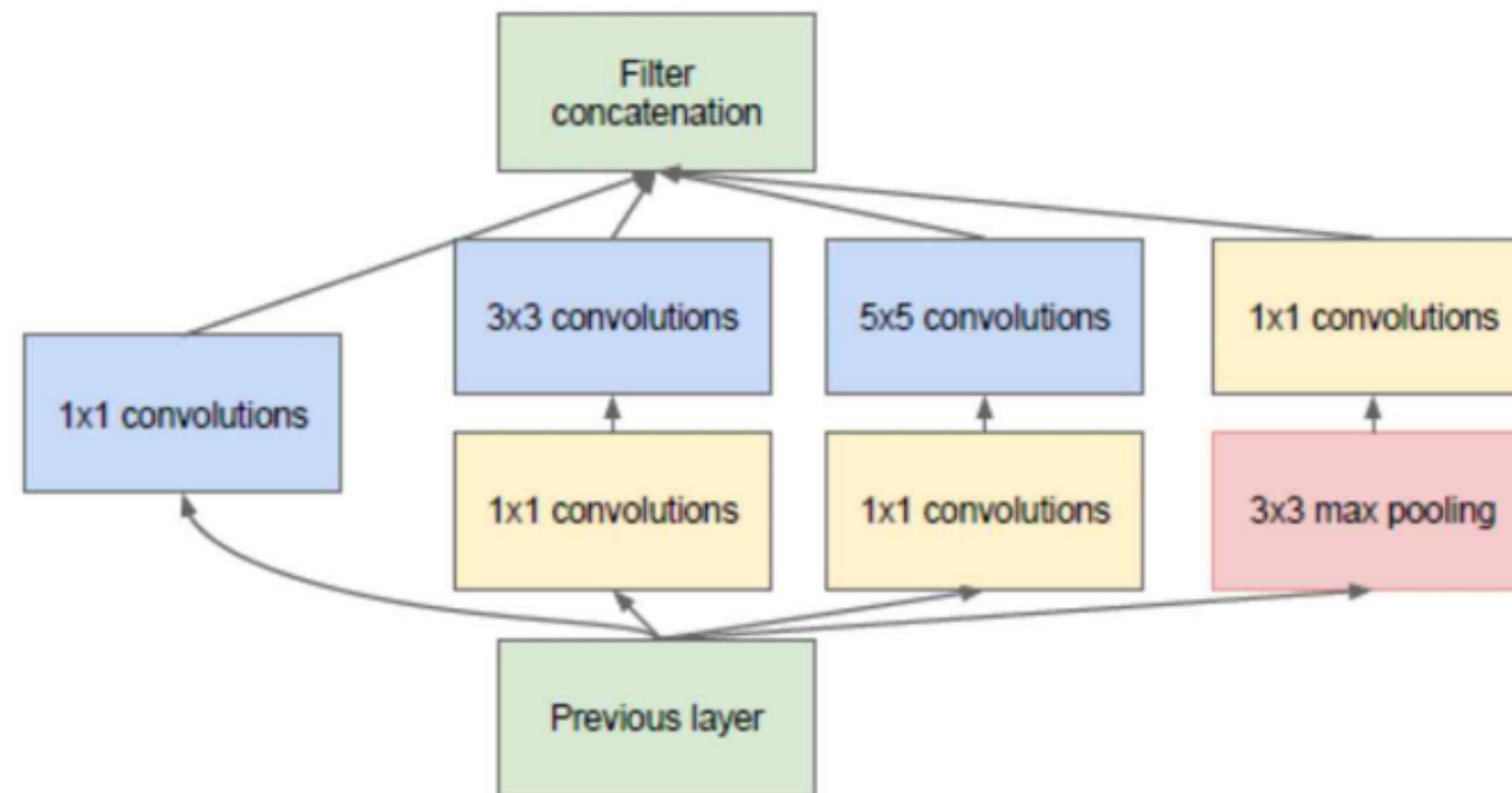
Layer 수가 16으로 증가 - 엄청난 성능 향상



Google Net

구조(1*1 conv 이용, Bottleneck...), 아이디어에 집중

세세한 구조는 Xception 리뷰할 기회가 있을 때 ... (skip)



Vanishing Gradient

Layer가 깊어질수록 좋은 성능을 낸다??

- 파라미터가 많아지고 과적합 위험
- 기울기 소실로 인한 학습의 어려움 ☆

20 layer vs 56 layer

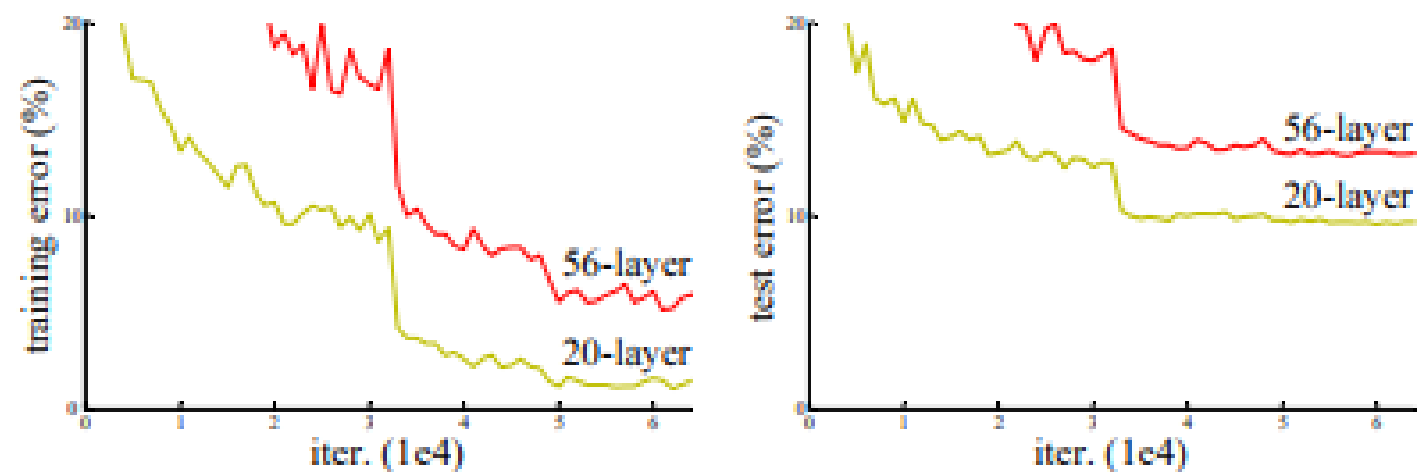
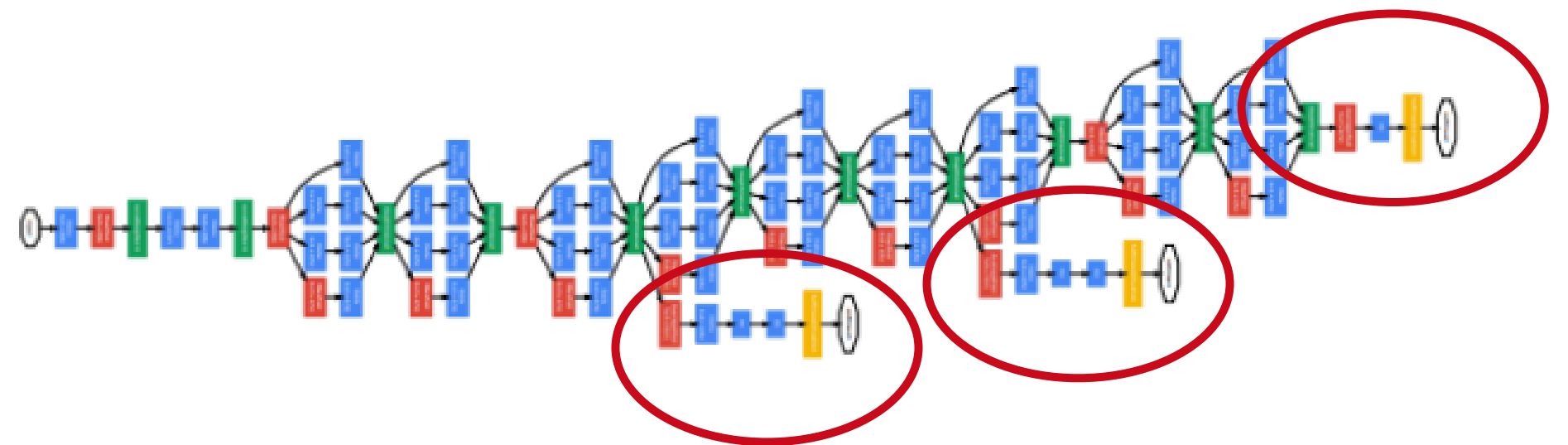


Figure 1. Training error (left) and test error (right) on CIFAR-10 with 20-layer and 56-layer “plain” networks. The deeper network has higher training error, and thus test error. Similar phenomena on ImageNet is presented in Fig. 4.

GoogLeNet의 기울기 소실 해결



보조의 Softmax Classifier 2개, Loss1&2
주 Softmax Classifier 1개, Original Loss
 $\text{Loss1} + 2 + \text{OgLoss} = \text{FinalLoss}$
각각 update → 정규화 효과

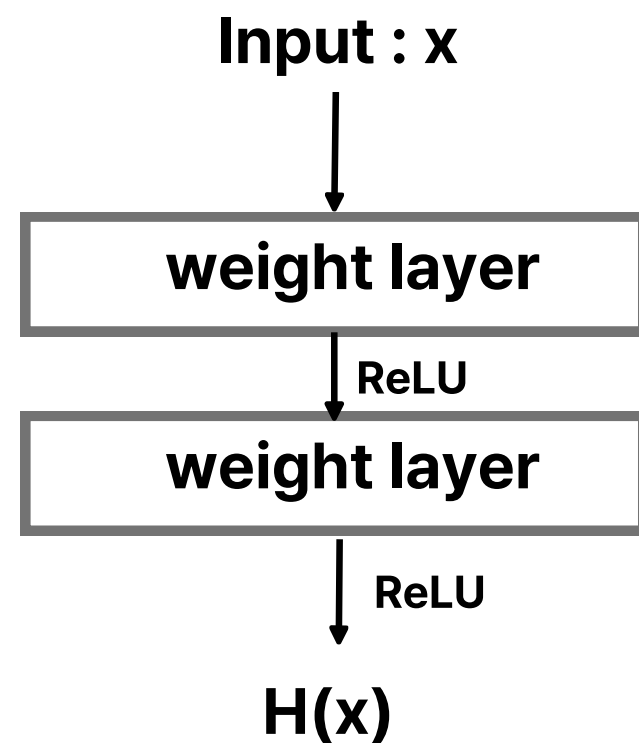
ResNet

Residual Learning

층이 깊어질수록 normalization, Weight Initialization 의 효과 떨어짐 -> Residual learning 주장

잔여정보 = $H(X)$ 와 $F(X)+X$ 가 같다는 가정(identity mapping), $H(X)$ 에서 기존의 정보 x 를 제외한 나머지 정보 = $F(X)$

shortcut을 추가함으로써 gradient가 사라진다고 해도 block를 학습하기 전의 x 를 추가로 더하기 때문에 기울기 손실 문제를 해결할 수 있다. ($F(X) + X$ 미분 -> $F(X) + 1$)



기존의 mapping

identity mapping



$$H(X) - X = F(X)$$

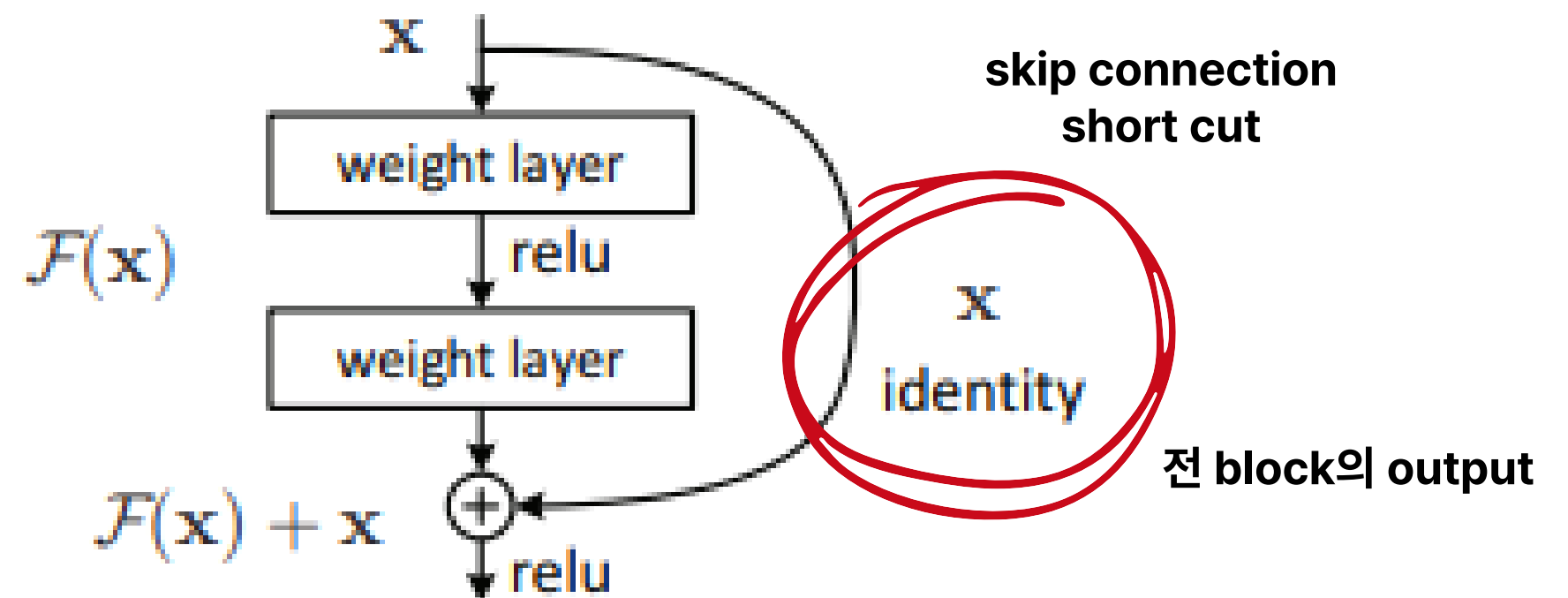


Figure 2. Residual learning: a building block.

Residual Learning

i가 1일 땐 기존의 identity mapping $H(X)$ 와 성능면에서 별 차이가 없기 때문에 weight layer가 웬만하면 2개 이상의 층을 갖는 것을 본 논문에선 권장한다.

$$y = \mathcal{F}(x, \{W_i\}) + W_s x.$$

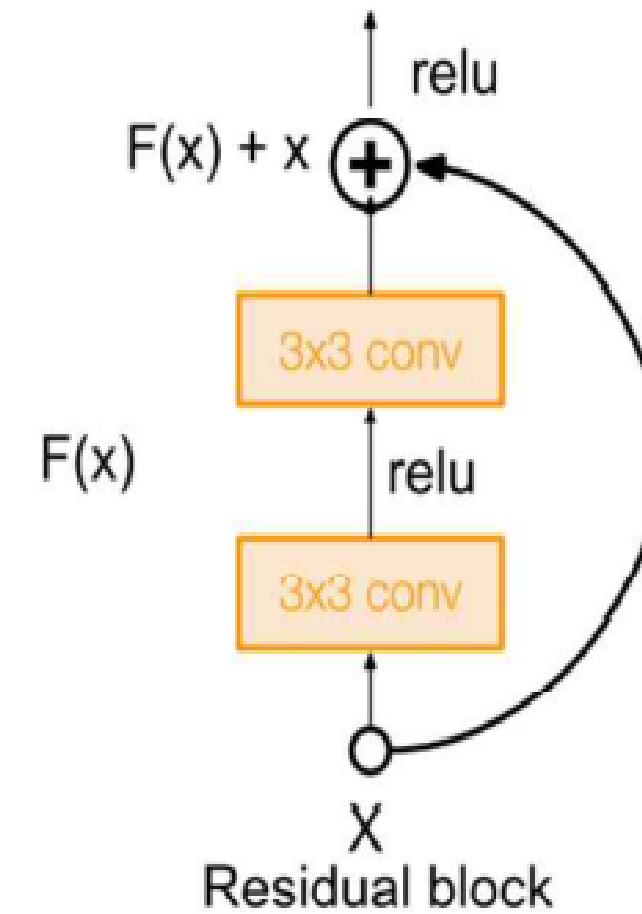
||

$$\mathcal{F} = W_2 \sigma(W_1 x)$$

x와 F는 서로 다른 차원을 가질 수 있음.
x를 F에 linear projection 하여 해결 가능
(1*1 convolution, padding...)

Residual Learning

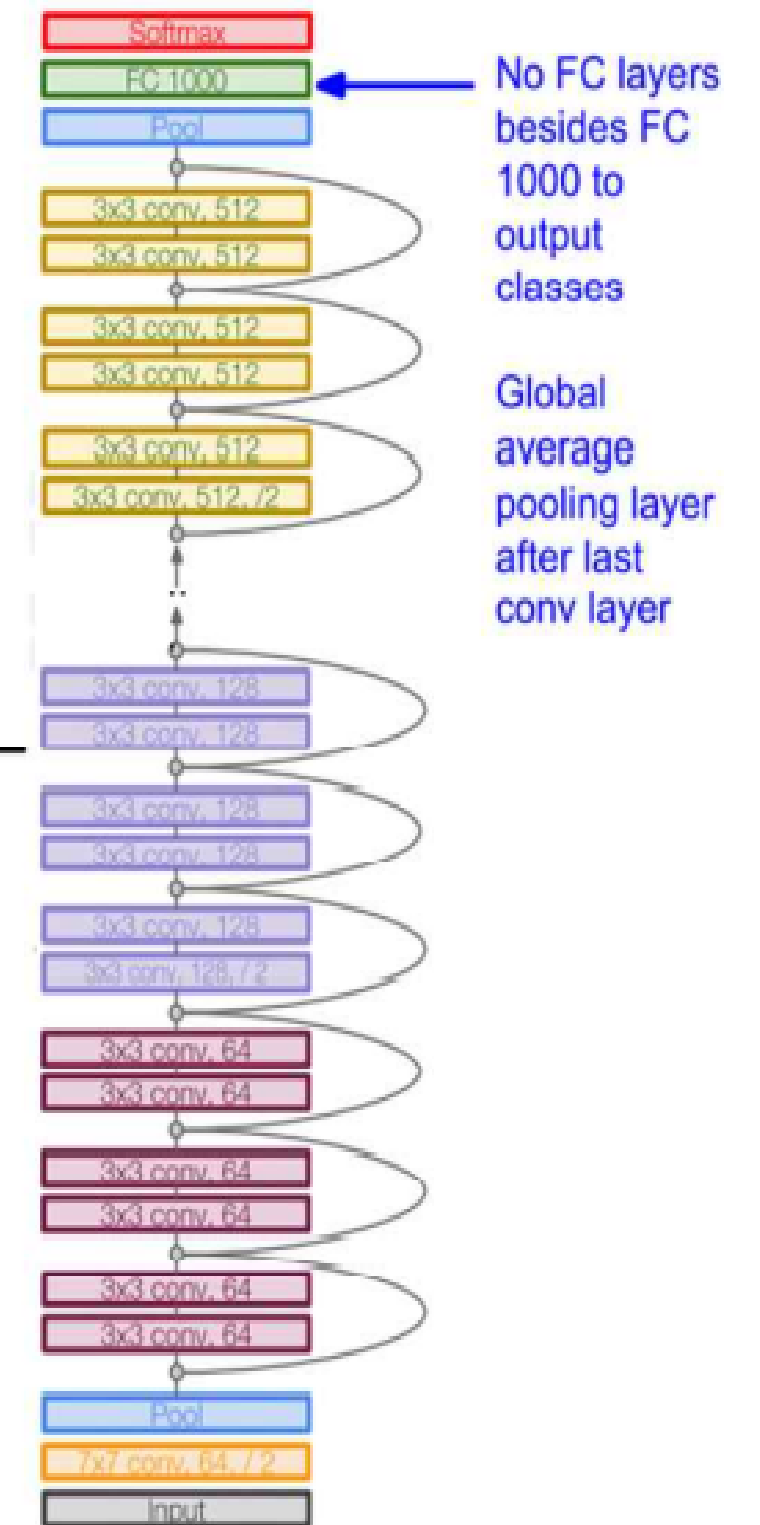
VGG구조에 Shortcut이 더해진 것이라고 봐도 된다.
Layer수가 기하급수적으로 많아져도 학습이 가능하다.



x
identity

Residual block

Fully-connected Layer를 1개만 남겨둠
Shortcut을 추가함

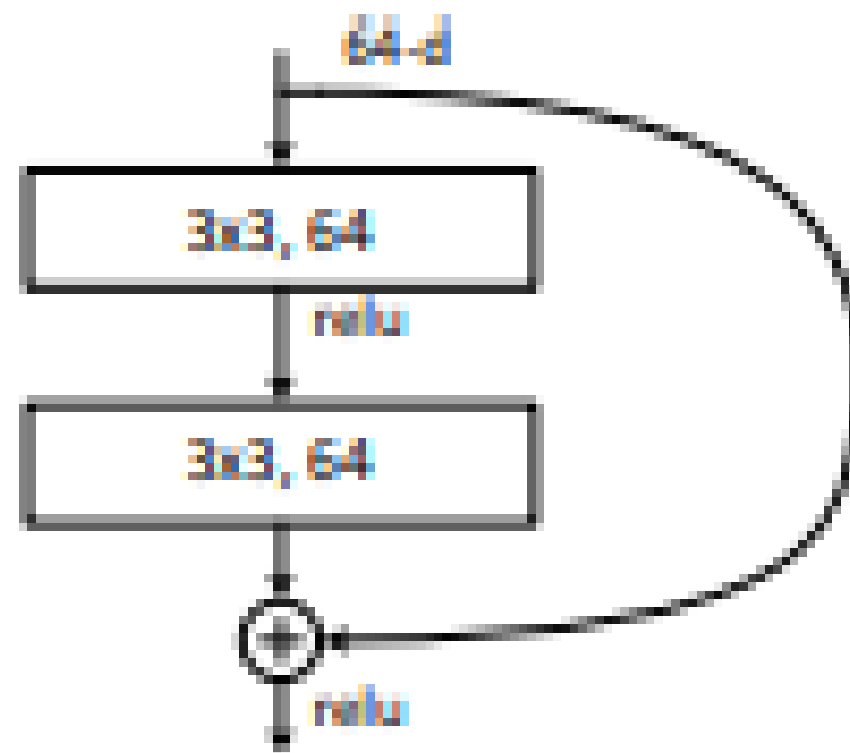


Bottle Neck

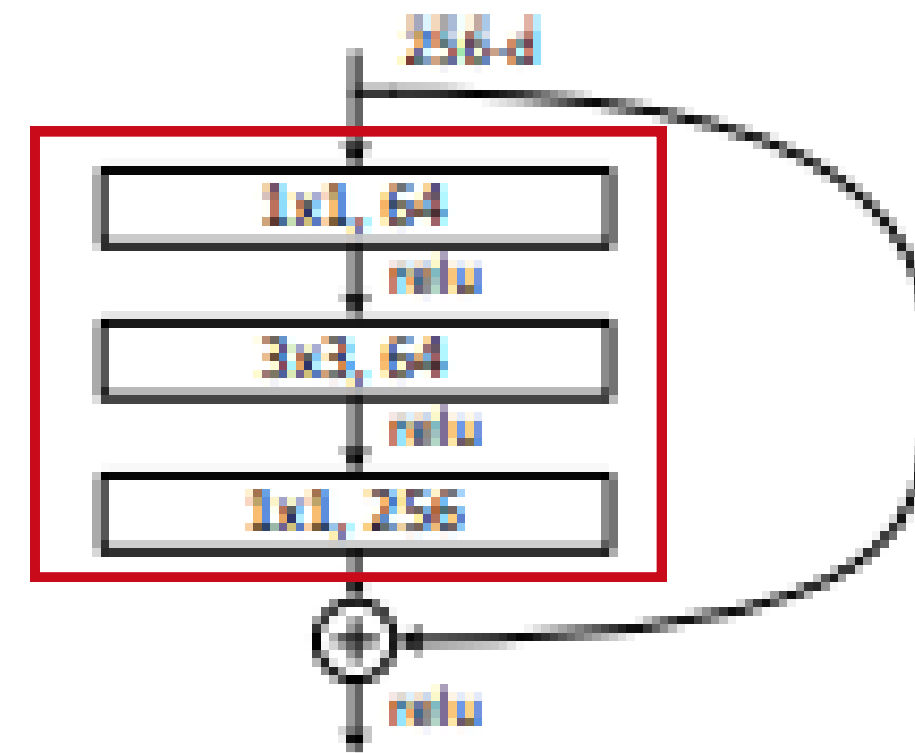
사회적 현상 중 병목현상과 닮았다하여 bottle neck 명명

1*1 conv (차원 축소) → 3*3 conv → 1*1 conv (차원 증가)

블록 내의 layer가 많아진다는 것 → activation function 증가 → data를 더 가공 가능



기존 building block

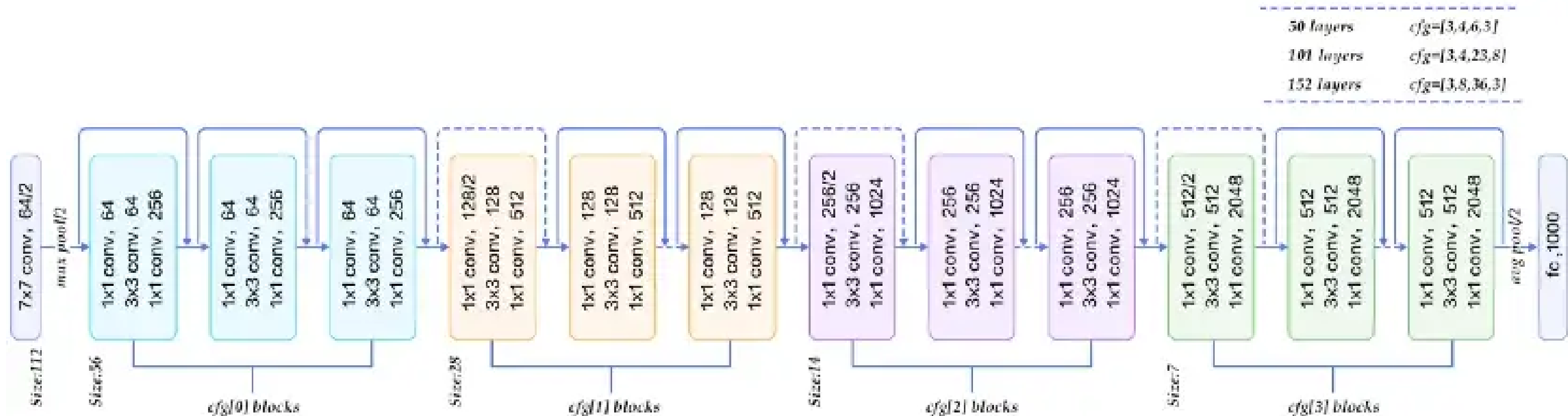


Bottle Neck

Bottle Neck

skip connection , bottle neck 으로 인해 더 많은 layer를 쌓을 수 있게 됨.

ResNet 50 부터는 bottle neck 구조로 아키텍처가 짜여짐

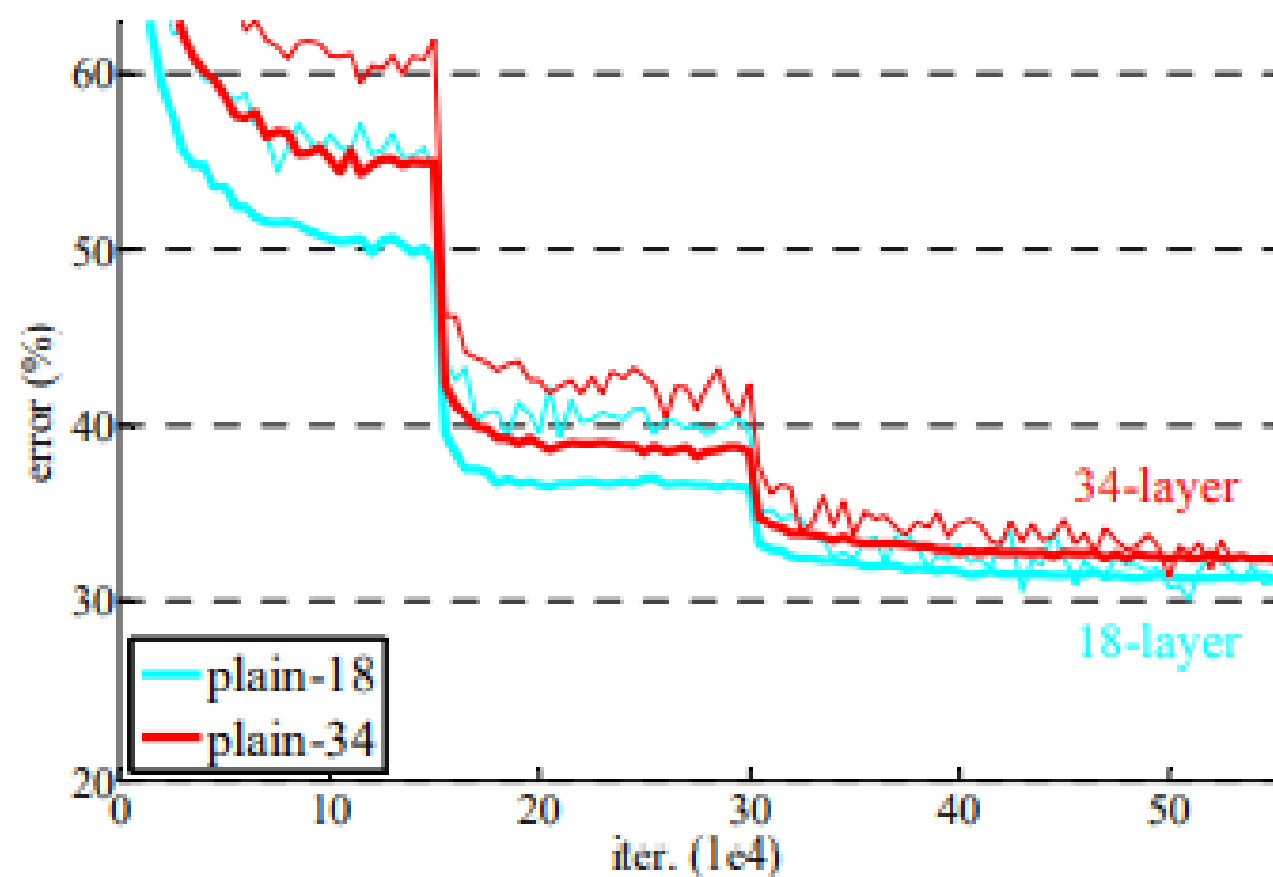


ImageNet 결과 분석

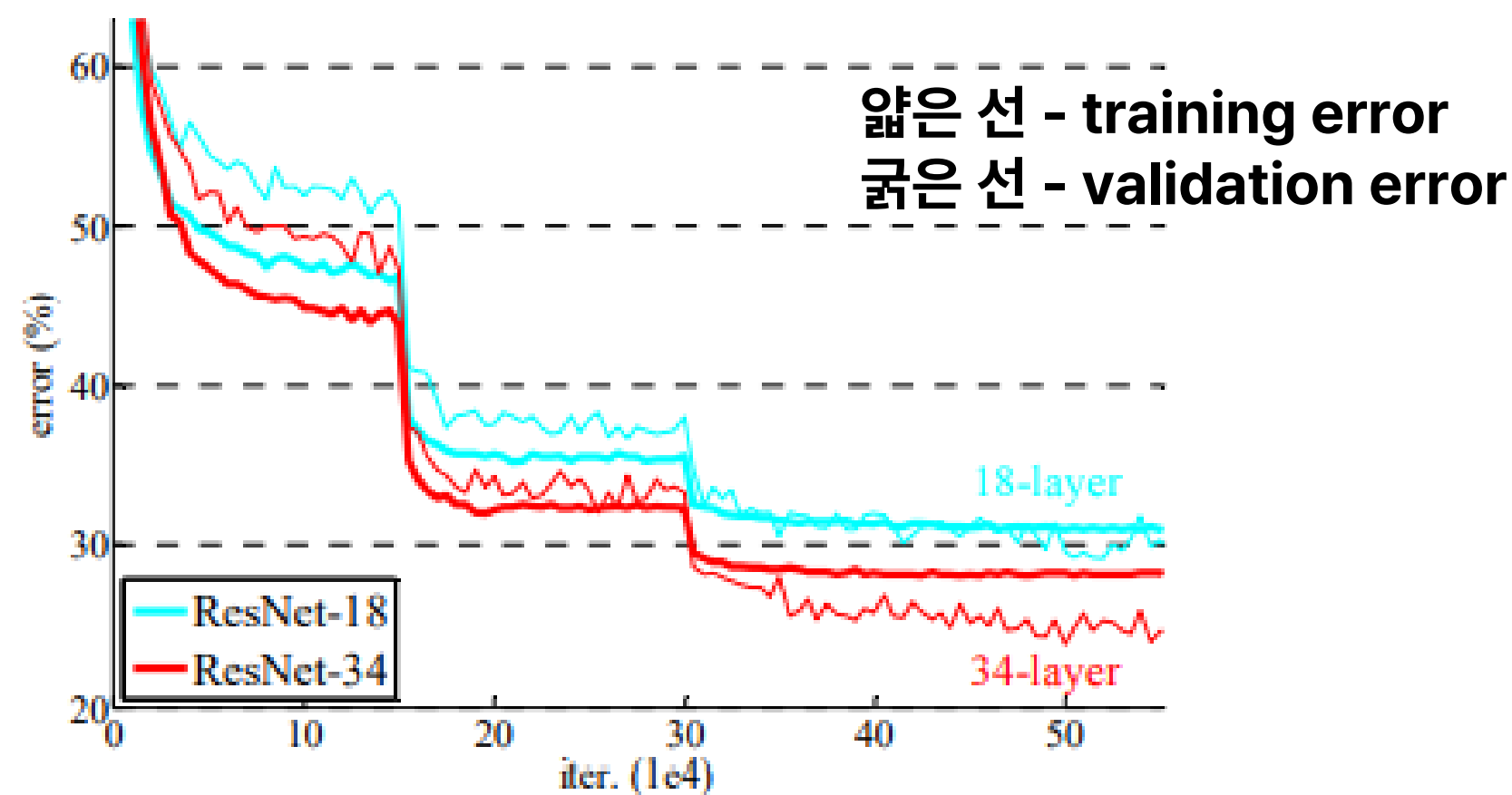
둘의 파라미터 수는 동일

기울기소실 문제 x
BN 사용, propagation 문제 안보임

	plain	ResNet
18 layers	27.94	27.88
34 layers	28.54	25.03



Plain CNN



ResNet

ImageNet 결과분석

A - Zero Padding
B - Projection Shortcut, identity
C - 모든 shortcut에 B적용

model	top-1 err.	top-5 err.
VGG-16 [41]	28.07	9.33
GoogLeNet [44]	-	9.15
PRelu-net [13]	24.27	7.38
plain-34	28.54	10.02
ResNet-34 A	25.03	7.76
ResNet-34 B	24.52	7.46
ResNet-34 C	24.19	7.40
ResNet-50	22.85	6.71
ResNet-101	21.75	6.05
ResNet-152	21.43	5.71

Image Net validation
Error rate

method	top-1 err.	top-5 err.
VGG [41] (ILSVRC'14)	-	8.43 [†]
GoogLeNet [44] (ILSVRC'14)	-	7.89
VGG [41] (v5)	24.4	7.1
PRelu-net [13]	21.59	5.71
BN-inception [16]	21.99	5.81
ResNet-34 B	21.84	5.71
ResNet-34 C	21.53	5.60
ResNet-50	20.74	5.25
ResNet-101	19.87	4.60
ResNet-152	19.38	4.49

단일 모델

method	top-5 err. (test)
VGG [41] (ILSVRC'14)	7.32
GoogLeNet [44] (ILSVRC'14)	6.66
VGG [41] (v5)	6.8
PRelu-net [13]	4.94
BN-inception [16]	4.82
ResNet (ILSVRC'15)	3.57

앙상블

감사합니다