

꼼꼼한 딥러닝 논문 리뷰와 코드 실습

Deep Learning Paper Review and Code Practice

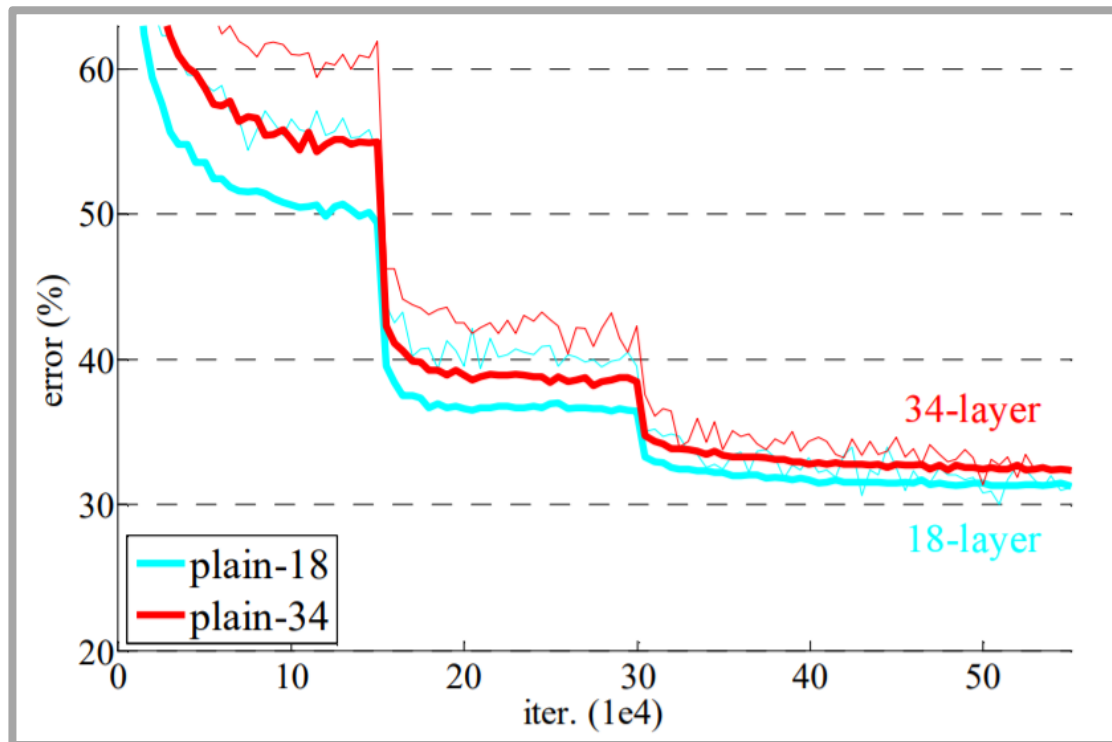
나동빈(dongbinna@postech.ac.kr)

Pohang University of Science and Technology

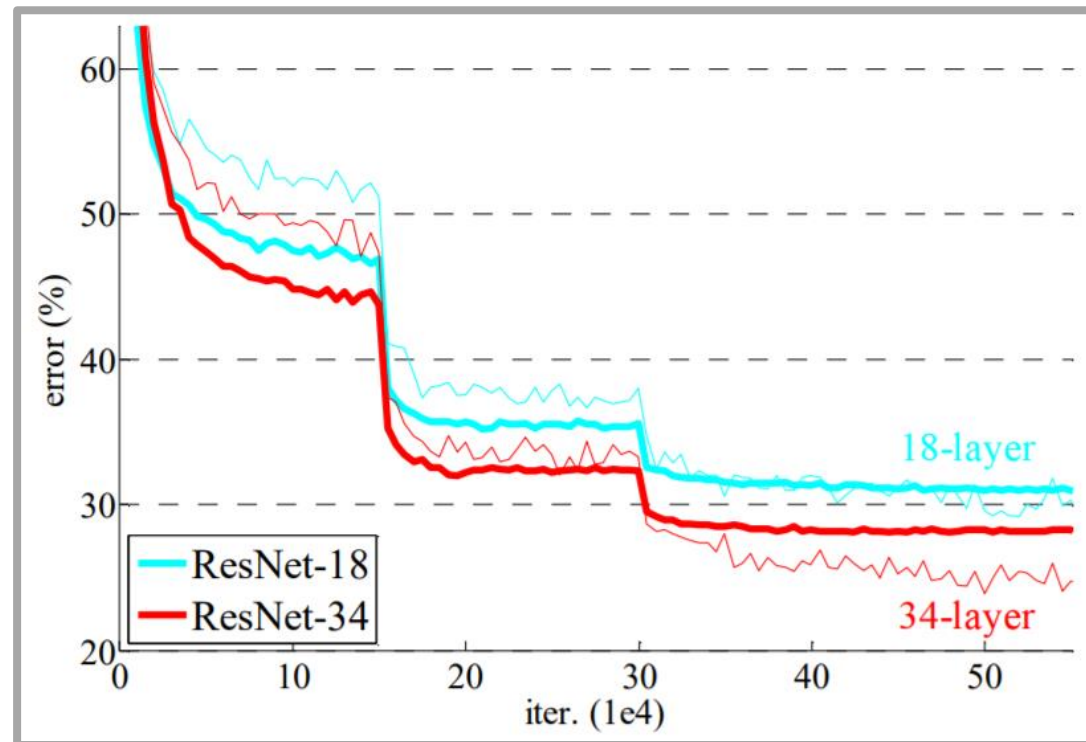
Deep Residual Learning for Image Recognition (CVPR 2016)

- 본 논문에서는 깊은 네트워크를 학습시키기 위한 방법으로 잔여 학습(residual learning)을 제안합니다.

< ImageNet top-1 training error >



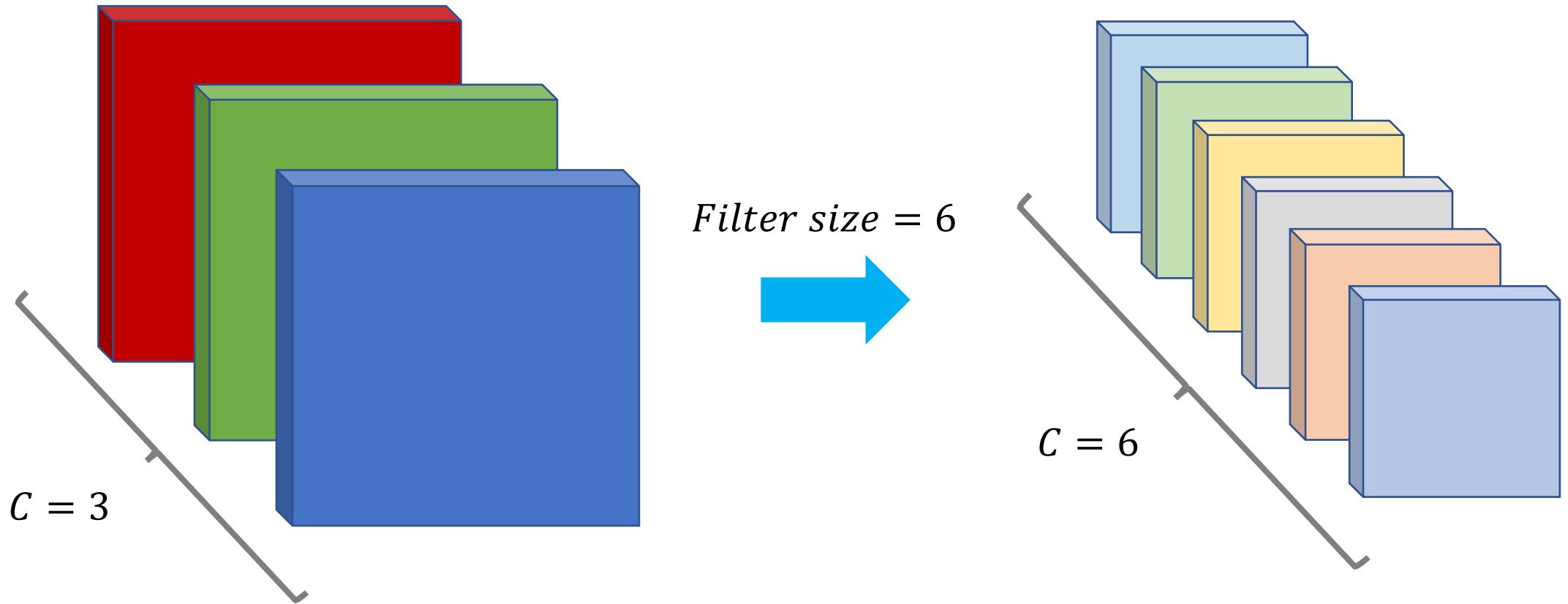
일반적인 CNN



잔여 학습을 적용한 CNN

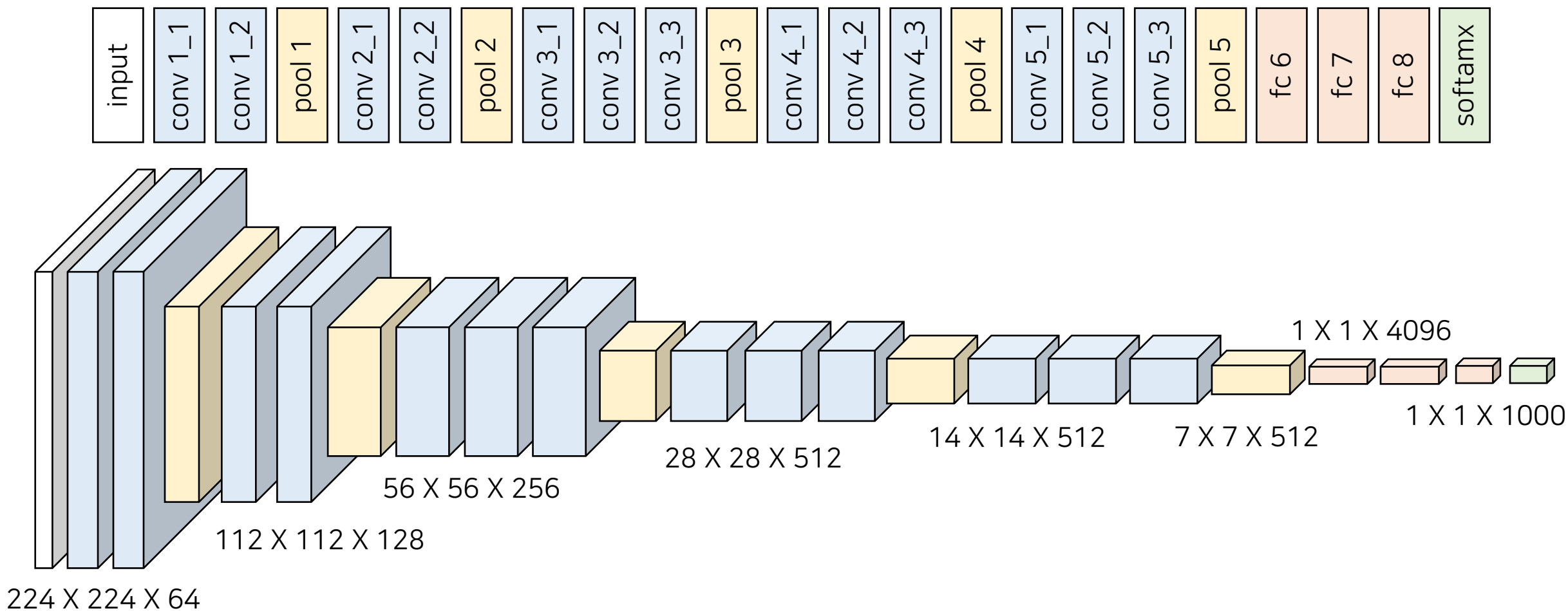
CNN 모델의 특징 맵 (Feature Map)

- 일반적으로 CNN에서 레이어가 깊어질수록 채널의 수가 많아지고 너비와 높이는 줄어듭니다.
- 컨볼루션 레이어의 서로 다른 필터들은 각각 적절한 특징(feature) 값을 추출하도록 학습됩니다.



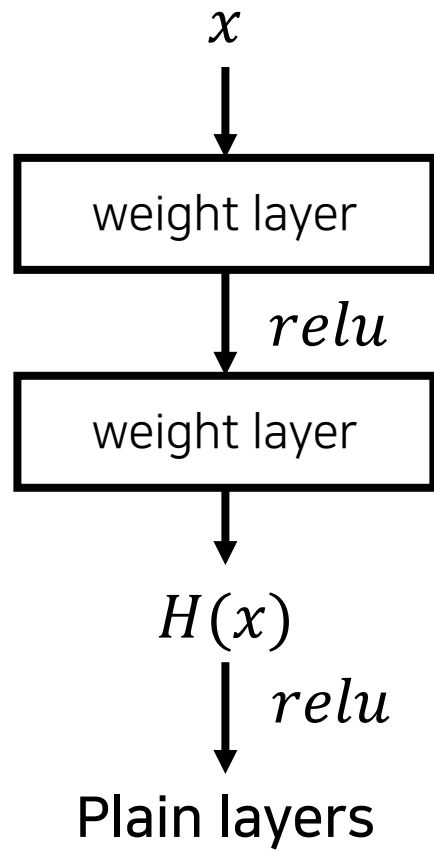
Very Deep Convolutional Networks for Large-Scale Image Recognition (ICLR 2015)

- VGG 네트워크는 작은 크기의 3x3 컨볼루션 필터(filter)를 이용해 레이어의 깊이를 늘려 우수한 성능을 보입니다.

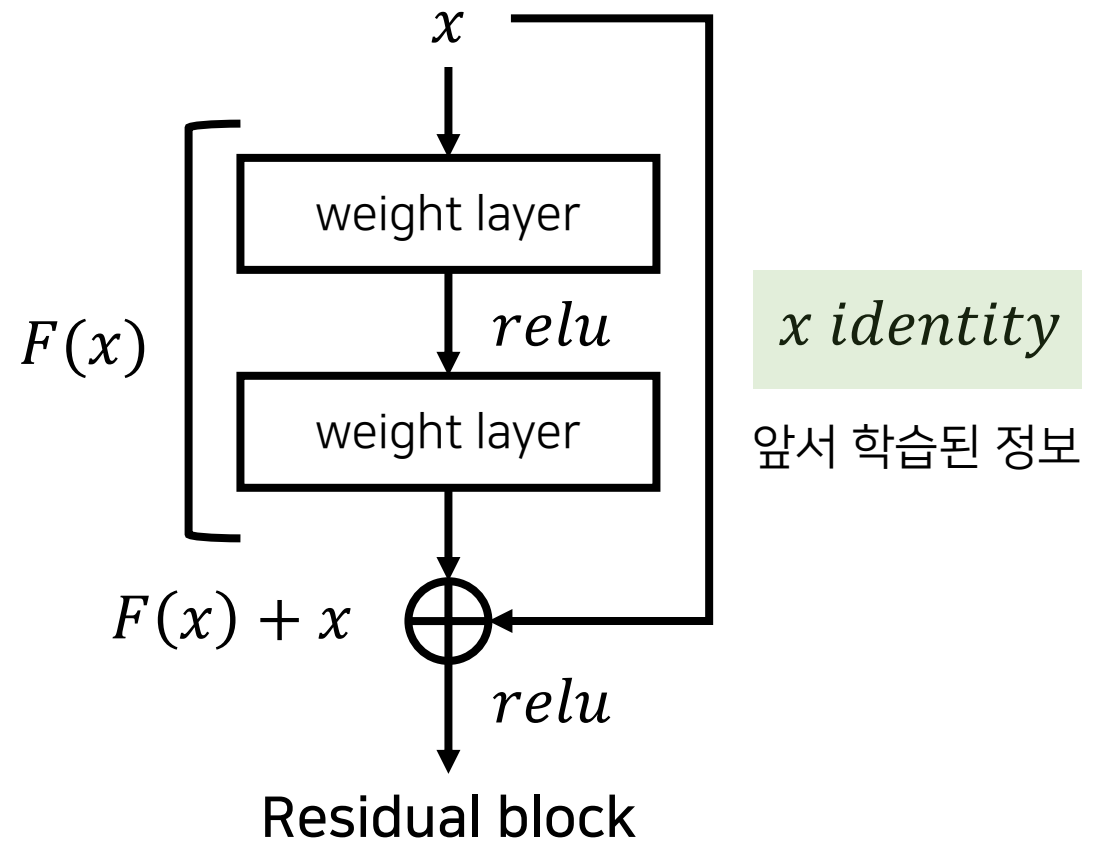


본 논문의 핵심 아이디어: 잔여 블록 (Residual Block)

- 잔여 블록(residual block)을 이용해 네트워크의 최적화(optimization) 난이도를 낮춥니다.
 - 실제로 내재한 mapping인 $H(x)$ 를 곧바로 학습하는 것은 어려우므로 대신 $F(x) = H(x) - x$ 를 학습합니다.

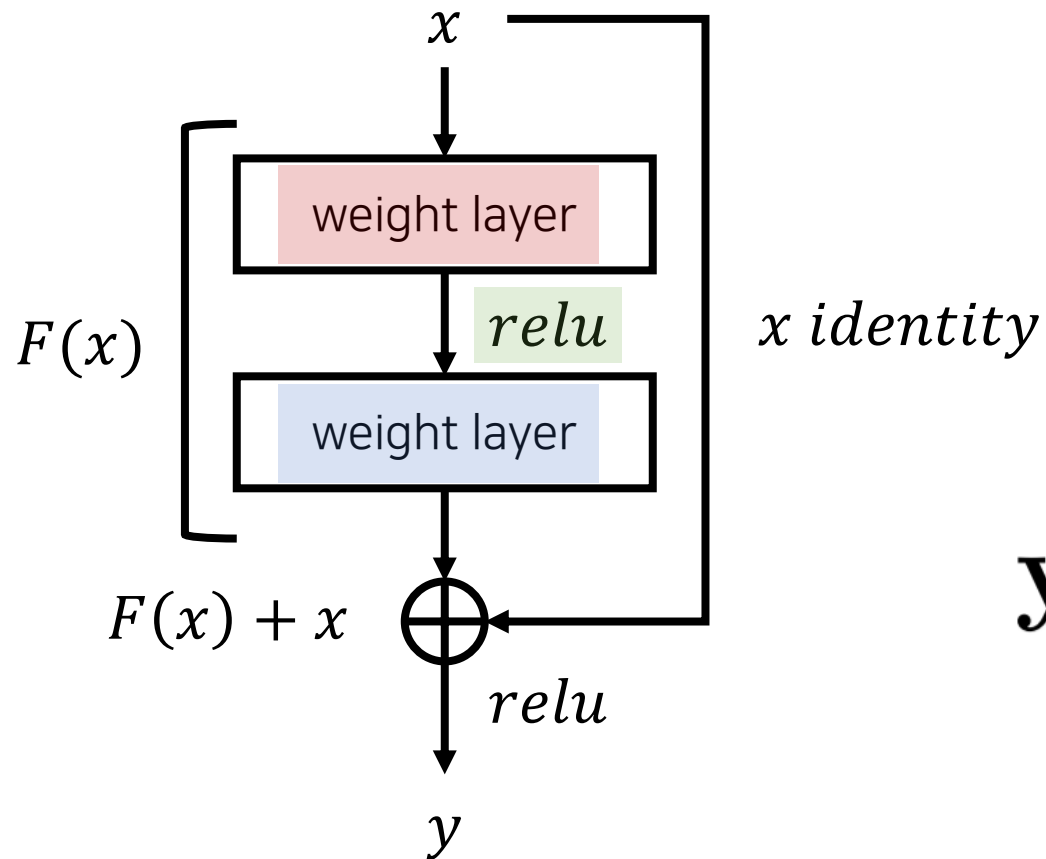


학습이 잘 되는 형태로 변경

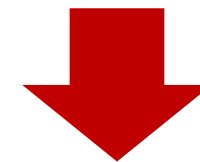


본 논문의 핵심 아이디어: 잔여 블록 (Residual Block)

- 잔여 블록(residual block)을 이용해 네트워크의 최적화(optimization) 난이도를 낮춥니다.



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



일반적인 형태

$$y = \underbrace{\mathcal{F}(\mathbf{x}, \{W_i\})}_{\text{multiple convolutional layers}} + \underbrace{W_s \mathbf{x}}_{\text{shortcut}}$$

ImageNet에서의 테스트 결과 분석

- 이전까지의 아키텍처와 다르게 레이어가 깊어질수록 성능이 향상됩니다. (단, 레이어가 과도하게 깊으면 오히려 감소)

	plain	ResNet
18 layers	27.94	27.88
34 layers	28.54	25.03

Top-1 validation error rates (%)

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

Top-5 test error rates (%) of ensembles

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

Validation error rates (%) of single-model