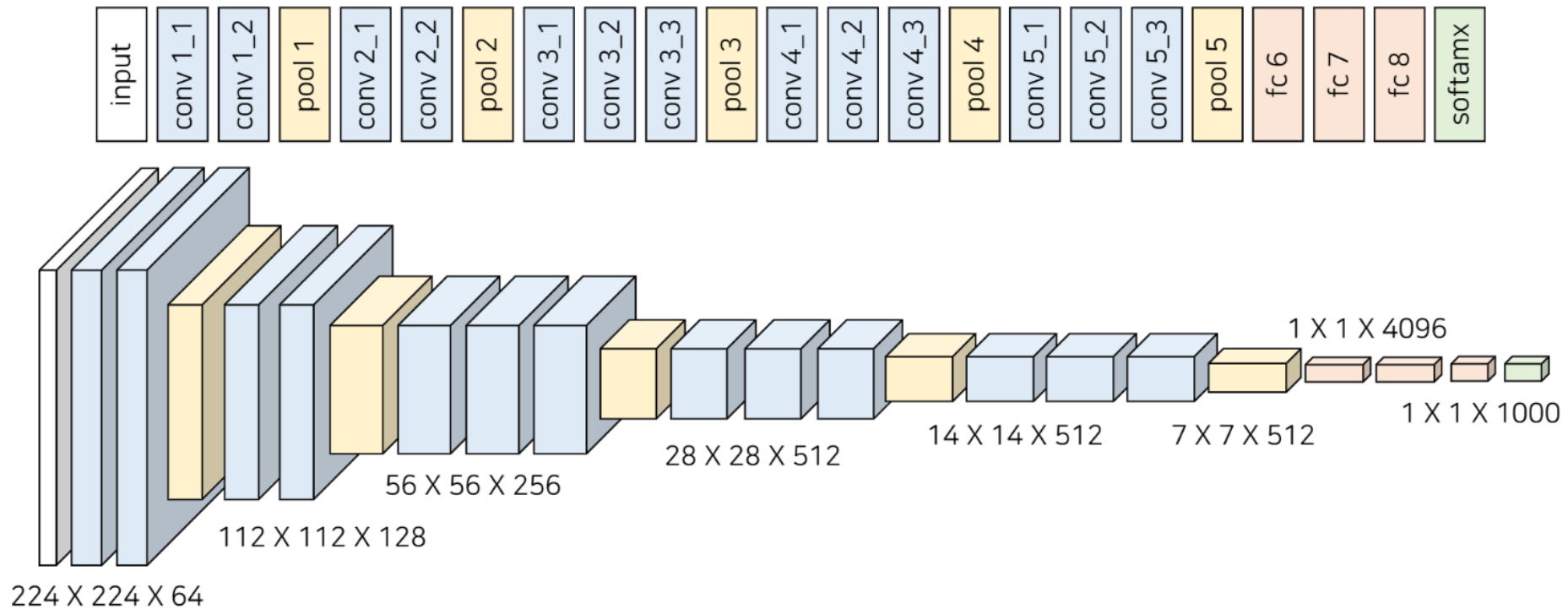


Deep Residual Learning for Image Recognition

:ResNet

CNN 모델의 Feature Map

- CNN에서 레이어가 깊어질수록 채널의 수가 많아지고, 너비와 높이는 줄어든다.
- Convolution layer의 서로 다른 필터들은 각각 적절한 feature 값을 추출하도록 학습된다.
- VGG네트워크 - 작은 크기의 3x3 convolution filter를 사용해 레이어의 깊이를 늘려 우수한 성능을 보임.



심층 신경망의 학습 어려움

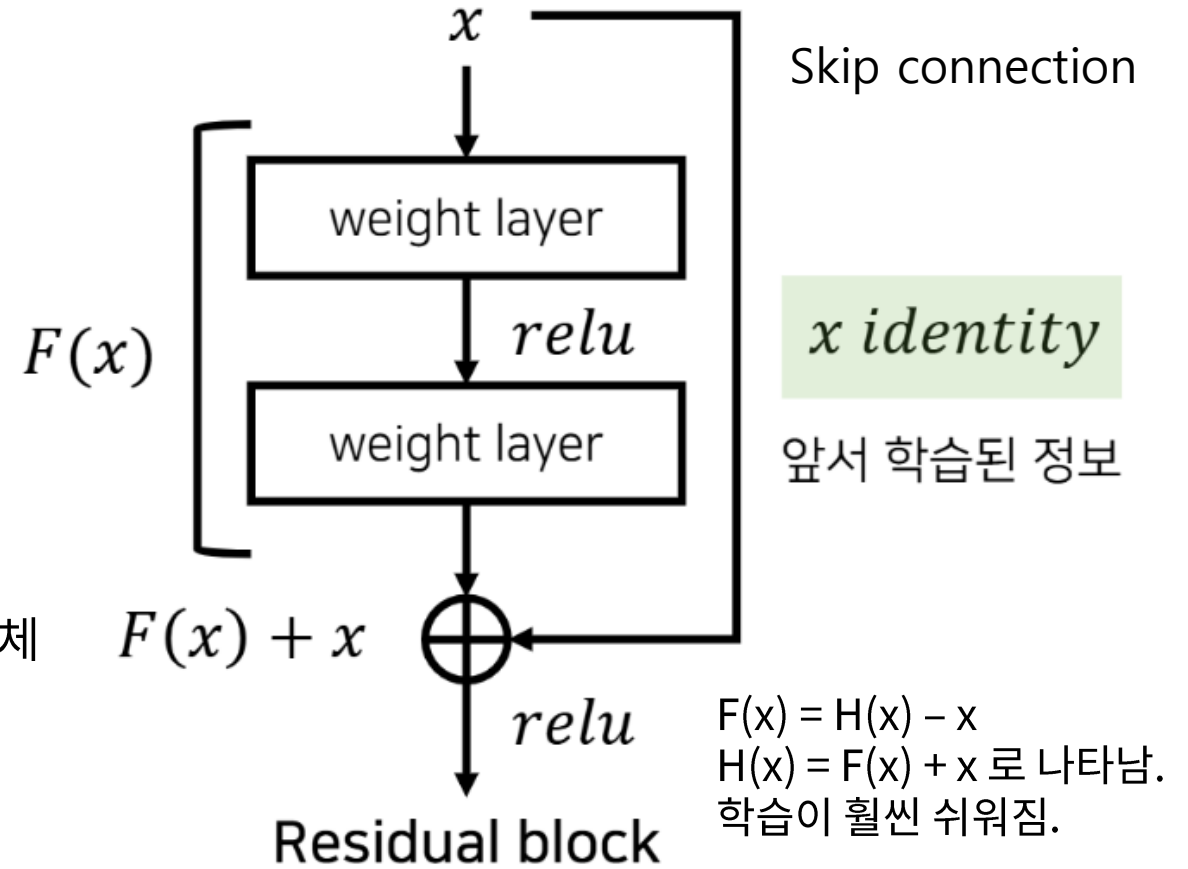
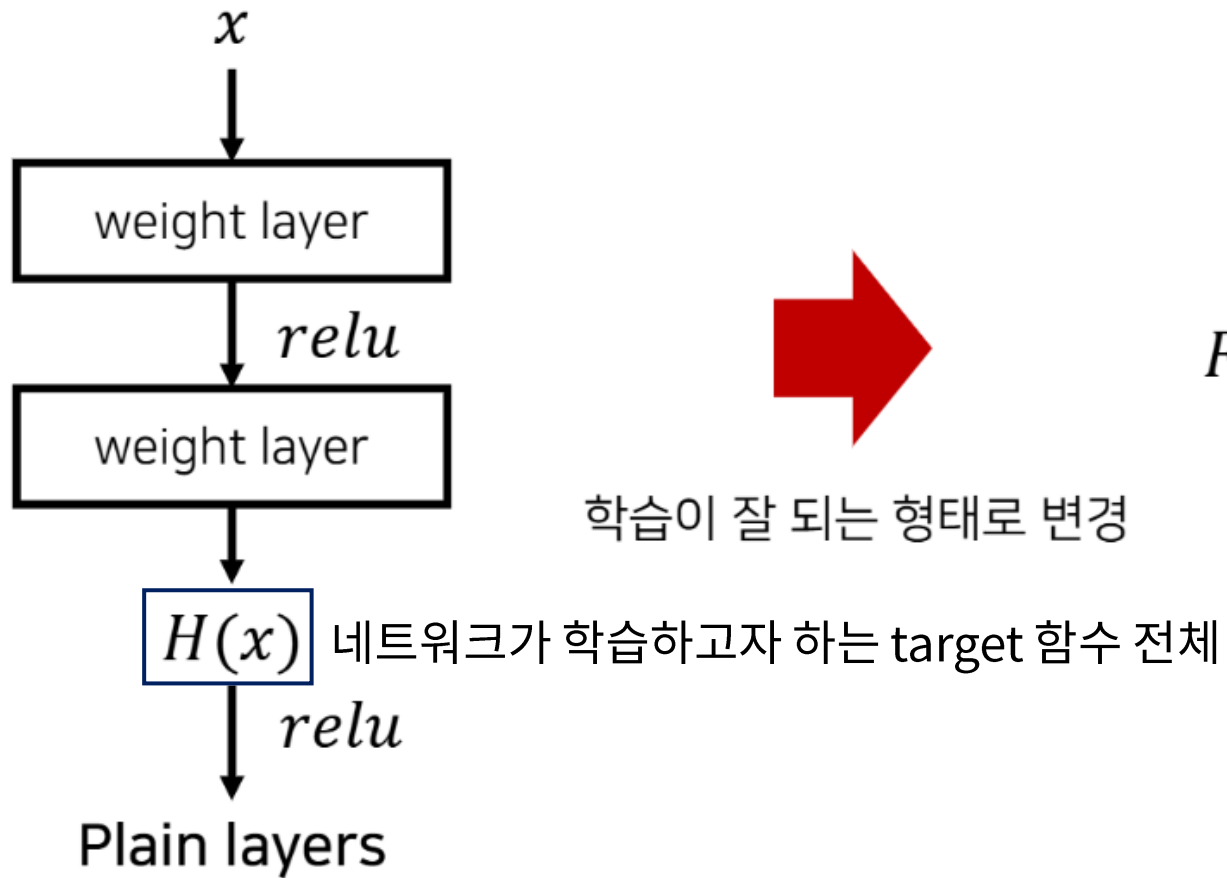
딥러닝에서 네트워크가
깊어질수록



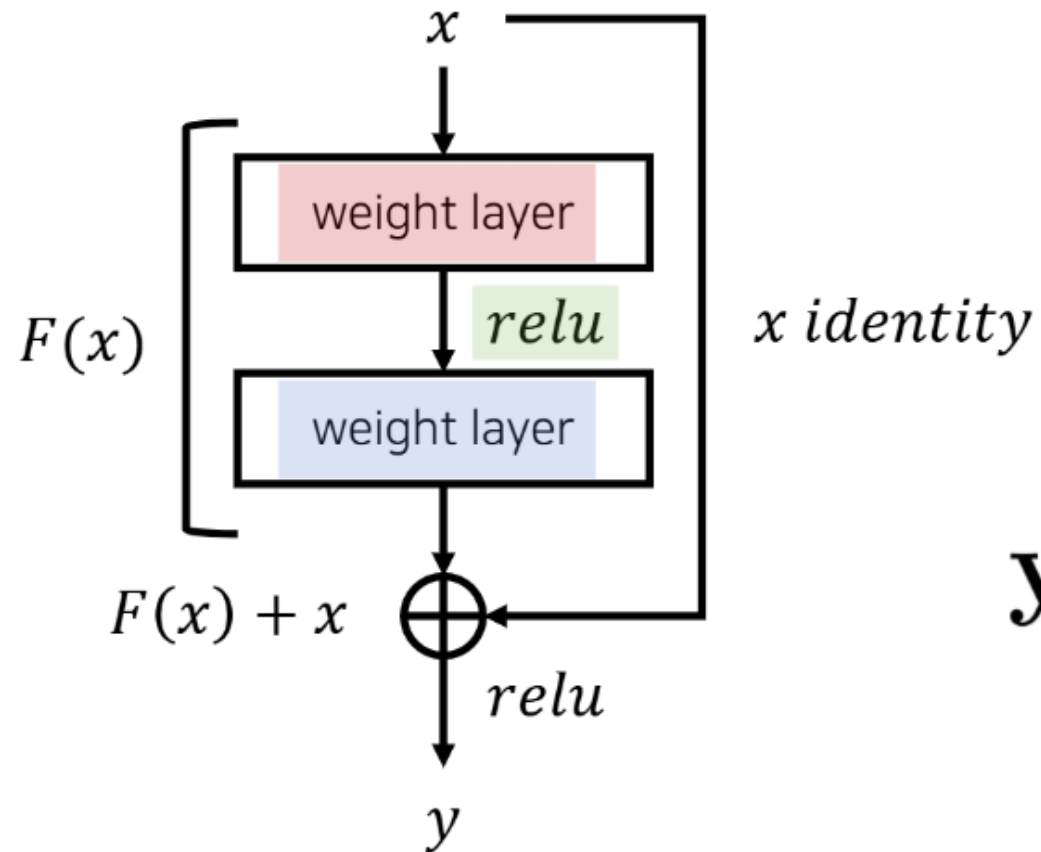
Vanishing Gradient
Exploding Gradient
성능 저하

Deeper neural networks를 학습시키기 위한 방법으로 residual learning 제안

Residual Block



Residual Block



$$\mathcal{F} = W_2 \overset{\text{Activation func}}{\sigma}(W_1 \mathbf{x})$$



일반적인 형태

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

Residual Block

- Identity Mapping: 입력 x 와 출력 y 가 동일한 dimension일 때
 - $F(x)$ 가 변형된 정보가 아니라, 입력 x 와 동일한 형태를 가지게 됨. $+x$
- Linear Projection: 입력 x 와 출력 y 가 다른 dimension일 때
 - Linear Transformation을 통해 차원을 맞춘다.
 - X 의 차원을 y 의 차원에 맞게 변환한다.
 - W_s : shortcut 경로의 weight 의미.
 - Ex. X 의 $d(256)$, y 의 $d(512)$ 일 때, 1×1 합성곱 (채널 수만 변환)을 사용하여 확장할 수 있음.

ImageNet에서의 Test 결과

	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

Insight

- 기존에는 더 깊은 네트워크를 사용해 학습하면 성능이 감소하는 문제가 있었음.
- Residual learning을 사용하면, 네트워크가 실제 학습하는 것은 residual function이므로 모델은 “residual”만 학습하면 됨.
- Shortcut connection은 overfitting을 유발하지 않음. (추가적인 파라미터가 거의 없기 때문에)
- CV -> Semantic segmentation에서 residual structure를 활용해 gradient 흐름 유지
- NLP -> Attention 기반 구조에서도 Layer Normalization 및 Feedforward layer에 Residual connection이 사용됨.
- => 거의 모든 딥러닝 영역에서 사용될만큼 범용성이 높음.