**(ResNet) Deep residual learning for image recognition**

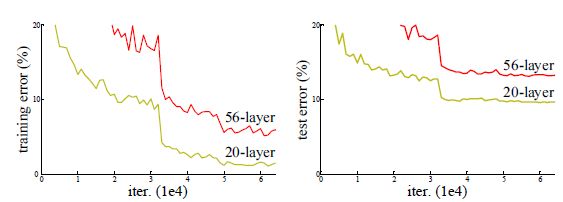
**2020.07.16**

**정경호**

**정병현**

* 개요

망이 깊어지면 성능이 좋아진다고 생각되지만 실제로는 성능 저하 문제가 드러난다. 네트워크의 depth가 증가하면서 정확도가 포화상태에 도달하게 되면 성능이 급격하게 떨어지는데 이를 degradation problem이라고 한다.



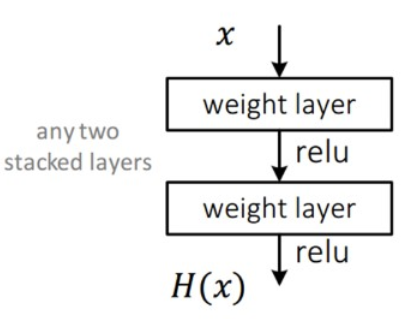
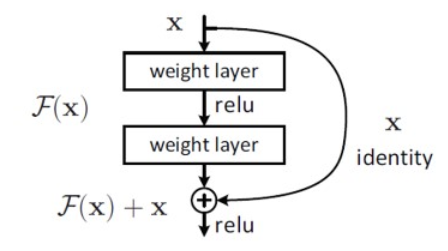
[그림 1]

그림 1은 20-layer와 56-layer를 가진 두 plain network가 CIFAR-10 dataset을 학습했을 떄의 training error(왼쪽)와 test error(오른쪽)를 나타낸다.

* ResNet

 original mapping을 H(x)로 나타낸다면, residual mapping은 H(x)=F(x)+x로 나타낸다. F(x)+x는 “shortcut connection”으로 구현할 수 있다. shortcut connection은 하나 이상의 layer를 건너 뛴다. shortcut connection은 identity mapping을 수행하고, 그 출력을 stacked layer의 출력에 더하고 있다.

H(x)=F(x)+x이기 때문에, H(x)=x가 되도록 한다는 것은 F(x)+x를 x가 되도록 학습한다는 뜻이다. 즉, residual(입력 x와의 잔차)인 F(x)가 0이 되게 하는 것이다.

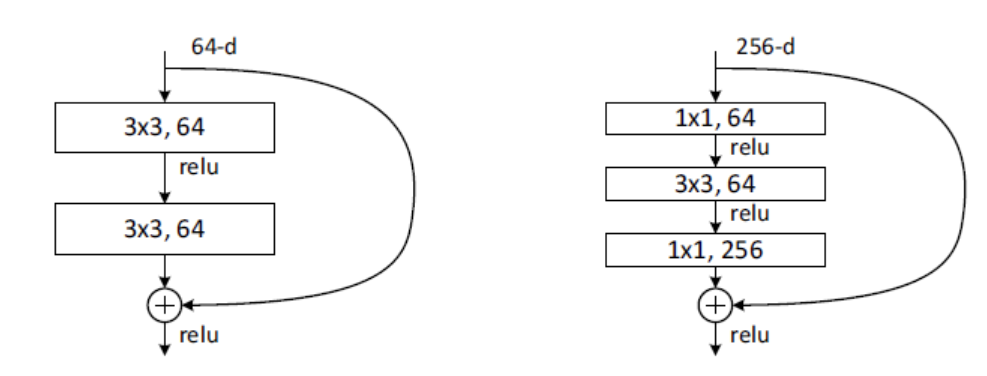
 

[그림 2]

그림 2는 입력 x가 2개의 stacked layer를 거친 결과(ReLU 이전)와 identity인 입력 x를 더한 후에 nonlinearity layer(ReLU)를 통과한다.

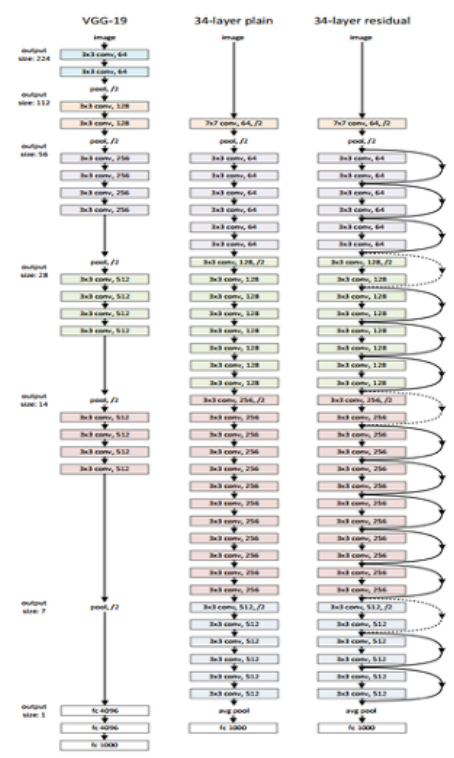
* Bottle Neck

34보다 깊은 망에서는 bottle neck구조를 사용한다. 차원이 큰 input을 처리할 때, 우선1x1 filter로 차원을 축소하고, 3x3 filter를 적용해서 feature를 추출한 뒤에 다시 1x1 filter로 차원을 원래대로 복구한다.

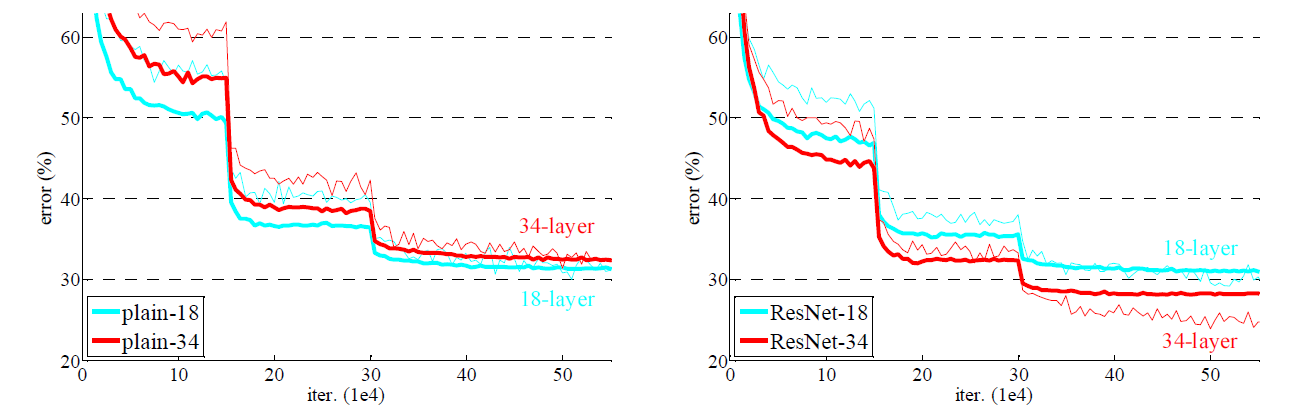


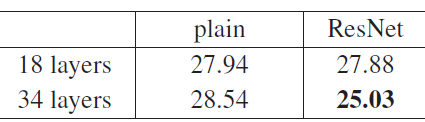
* 실험
* plain vs ResNet 비교

<Architecture>



<Result>

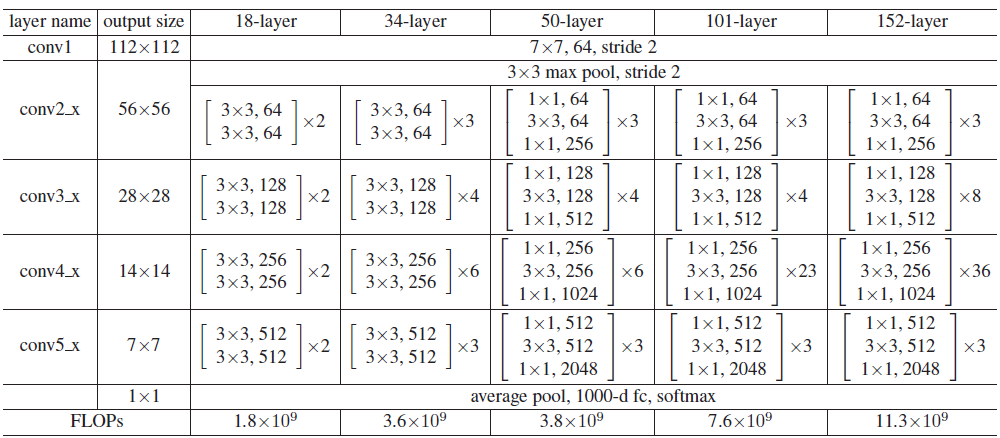




* 34, 50, 101, 152 layer ResNet 비교

ResNet-50, ResNet-101, ResNet-152은 2-layer block을 3-layer block으로 교체

< Architecture>



<Result>

