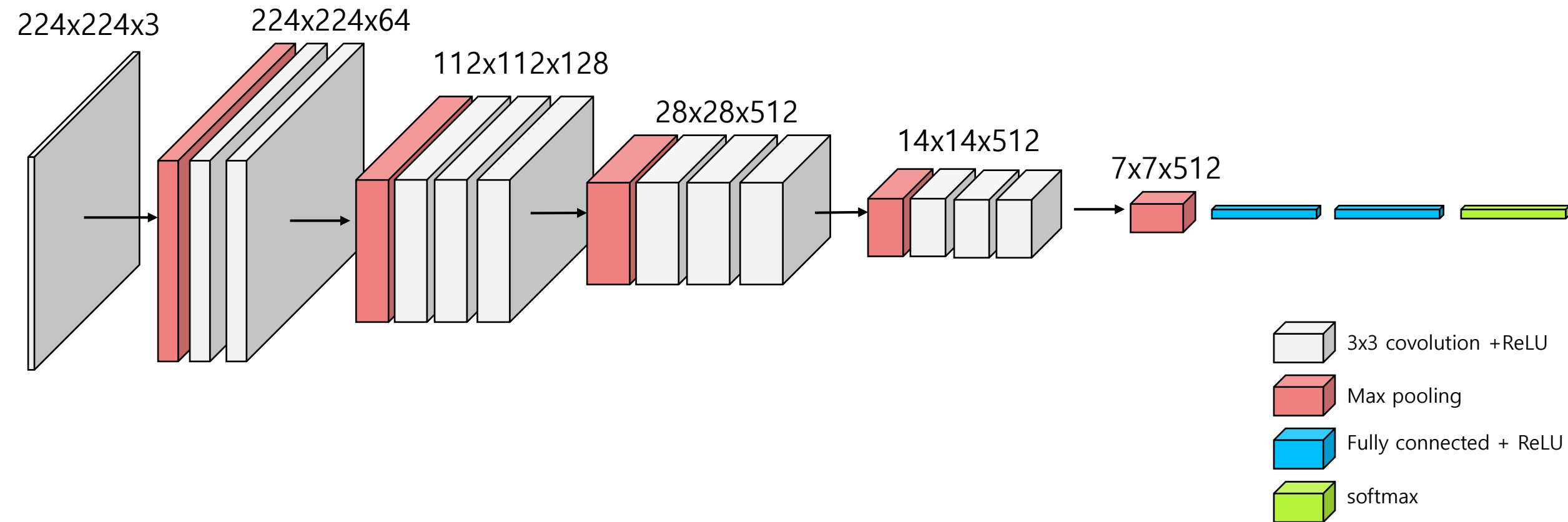


Deep Residual Learning for Image Recognition

Written in 2022.08.18

Research Background

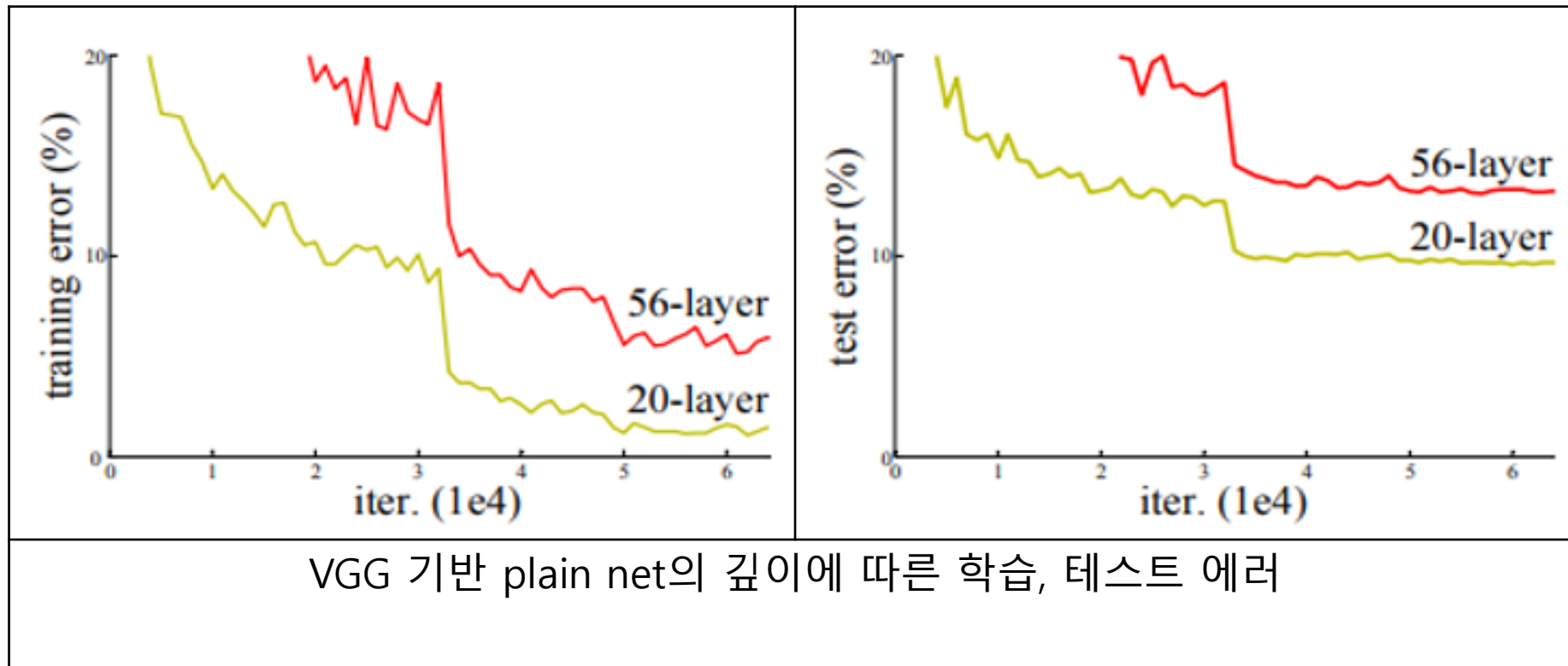
- 컴퓨터 비전 작업에서 레이어가 많이 쌓일 수록 좋은 성능을 보임
- VGG16은 작은 크기의 3x3 convolution 필터를 이용해 레이어의 깊이를 늘려 좋은 성능 보임



Research Background

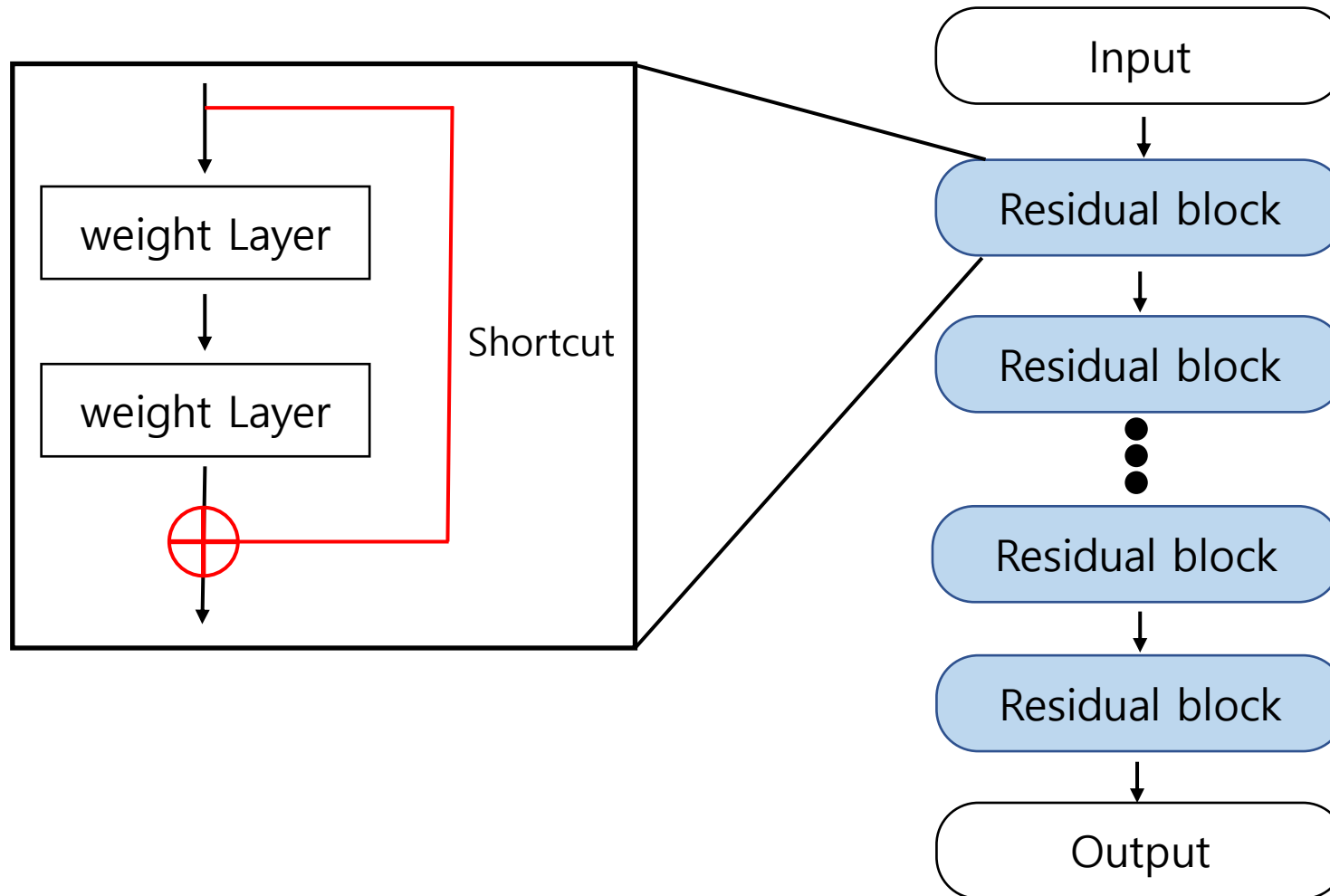
- 레이어가 많이 쌓이면 학습이 어려워지고, *Degradation 문제가 발생
- 이 문제를 해결하기 위해 Residual block을 제안

*Degradation: vanishing gradient/ exploding 과 다르게 학습 에러가 높을 때 테스트 에러 또한 높다.



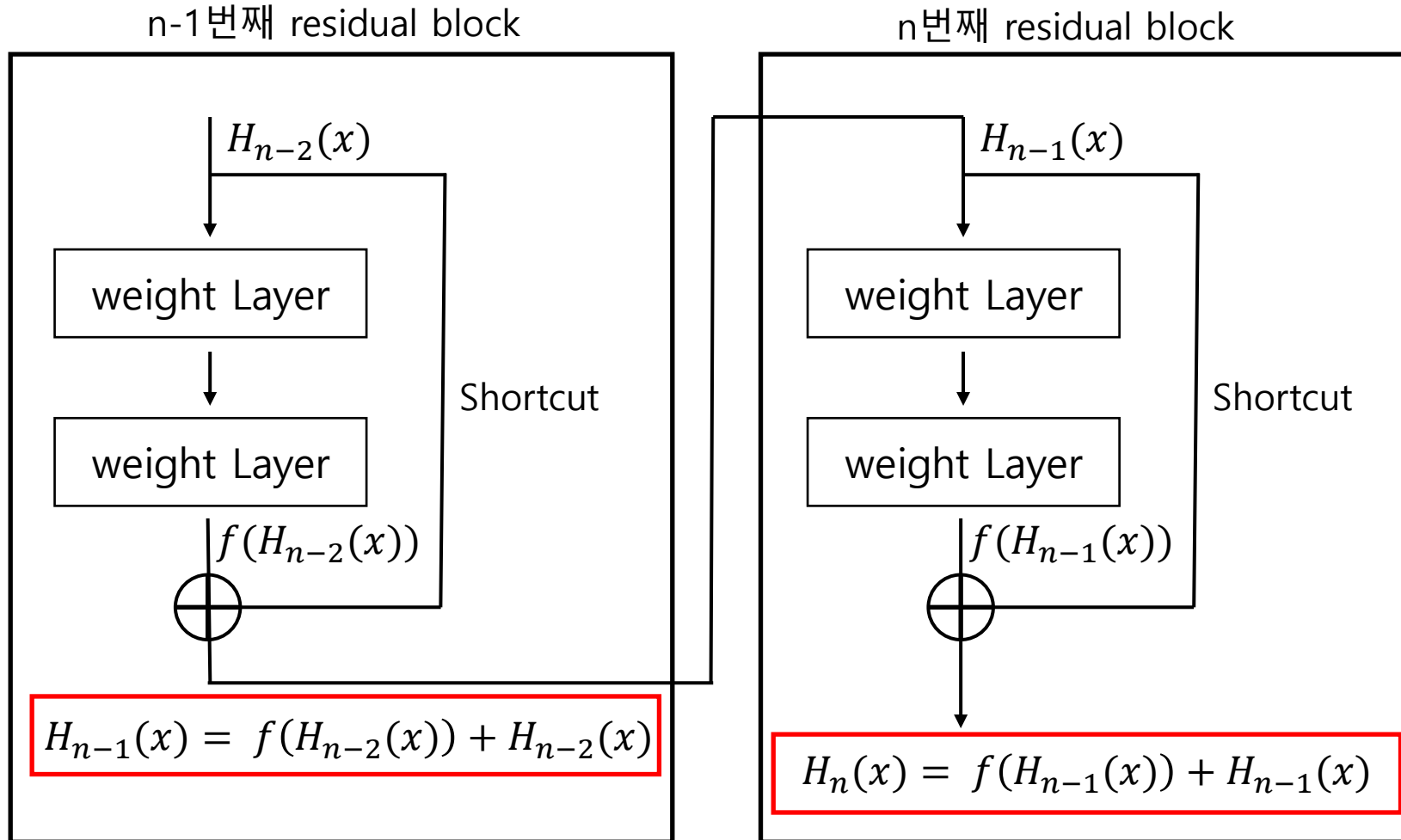
Model architecture

- Residual block은 입력 값과 출력 값을 더해주는 구조. 이를 위해 shortcut을 추가하여 입력과 출력을 연결
- 이 구조는 단순히 shortcut을 추가하여 더해주는 연산만 추가하여 파라미터 수가 일정.
- Residual network는 residual block을 여러 개 쌓은 구조



Mathematical expression

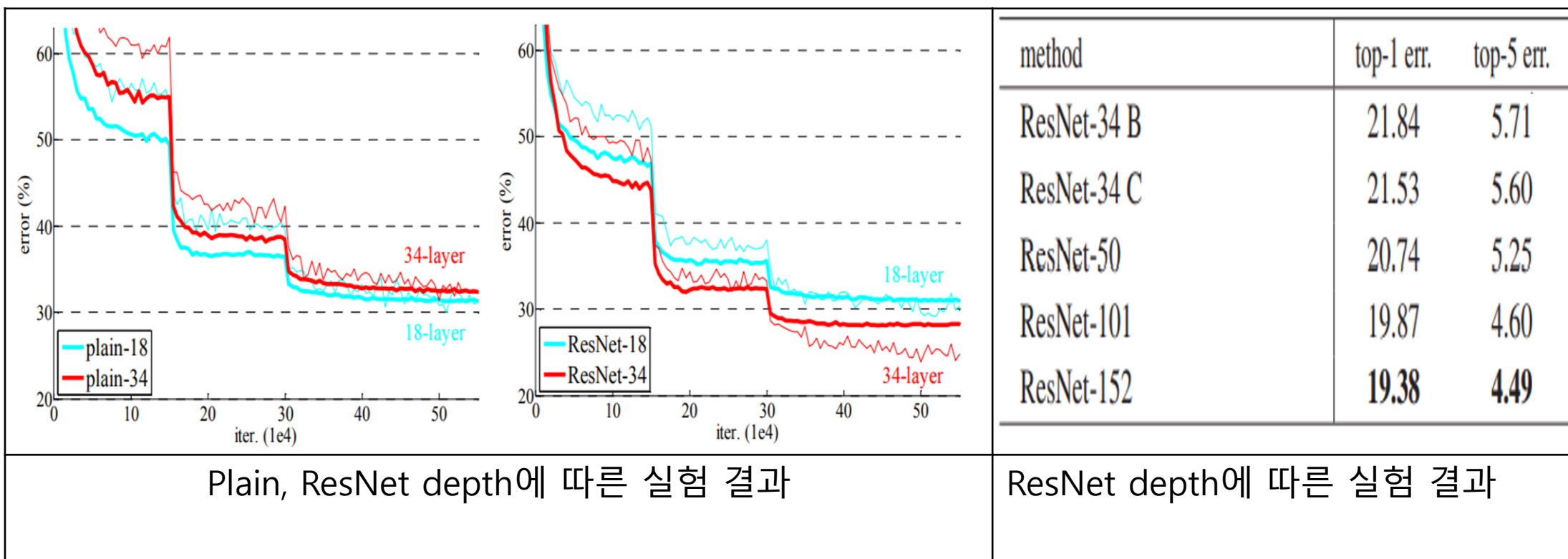
- Residual network의 출력은 $\sum_1^{n-1} f(H_k(x)) + x$ 이기에 가중치는 입력과 출력의 오차가 0이 되게 학습됨
- 그리고 VGG network와 다르게 덧셈 연산이 되고, 역전파할 때 최소 1 이상이 되기에 vanishing gradient가 발생하지 않음



$$\begin{aligned} H_n(x) &= f(H_{n-1}(x)) + H_{n-1}(x) \\ &= f(H_{n-1}(x)) + f(H_{n-2}(x)) + H_{n-2}(x) \\ &\quad \vdots \\ &= \sum_1^{n-1} f(H_k(x)) + x \\ - \quad &\frac{\partial \varepsilon}{\partial (H_n(x))} \left(\sum_1^{n-1} f(H_k(x)) + 1 \right) \end{aligned}$$

Result

- VGG기반 plain과 ResNet의 깊이를 증가시키며 학습과 테스트의 에러를 비교했을 때 degradation 문제가 해결됨.
- 152개의 레이어까지 쌓았을 때 degradation 문제가 발생하지 않고 더 좋은 성능을 보임



- 이전의 다른 모델들은 모델을 복잡하게 만들고, 새로운 방식들을 써서 약간의 층을 늘렸다면, 이 방식은 shortcut만을 추가하여 모델의 깊이를 압도적으로 깊게 만든 것이 인상적임