

CNN은 지금 이미지 인식 분야에서 선구자적인 역할을 맡고 있다.

CNN과 End – to –End Network가 합쳐져서 depth(stacked layer) 라는 단어까지 생겼다.

-> ImageNet Challenge를 통해 Network를 '아주 깊게' 쌓으면 기존보다 훨씬 좋은 성능이 나온다는 것을 확인하였다.

Q) 과연 Network의 Depth만 늘리는 것만으로도 성능이 쉽게 향상시킬 수 있는 것인가?

A) 그렇지 않다. 실제로는 Depth가 어느 부분까지 상승하다가 그 이상은 증가하지 않는다.
vanishing / exploding gradient 라는 문제 때문에 오히려 성능이 떨어지는 현상이 발생한다.
-> 이 현상을 **degradation problem**이라고 한다.
-> depth가 깊은 상태에서 학습을 많이 진행한 경우 weight들의 분포가 균등하지 않고,
backpropagation 시 gradient가 거의 0으로 수렴하여 학습을 안정적으로 진행할 수 없는 문제가 발생한다.

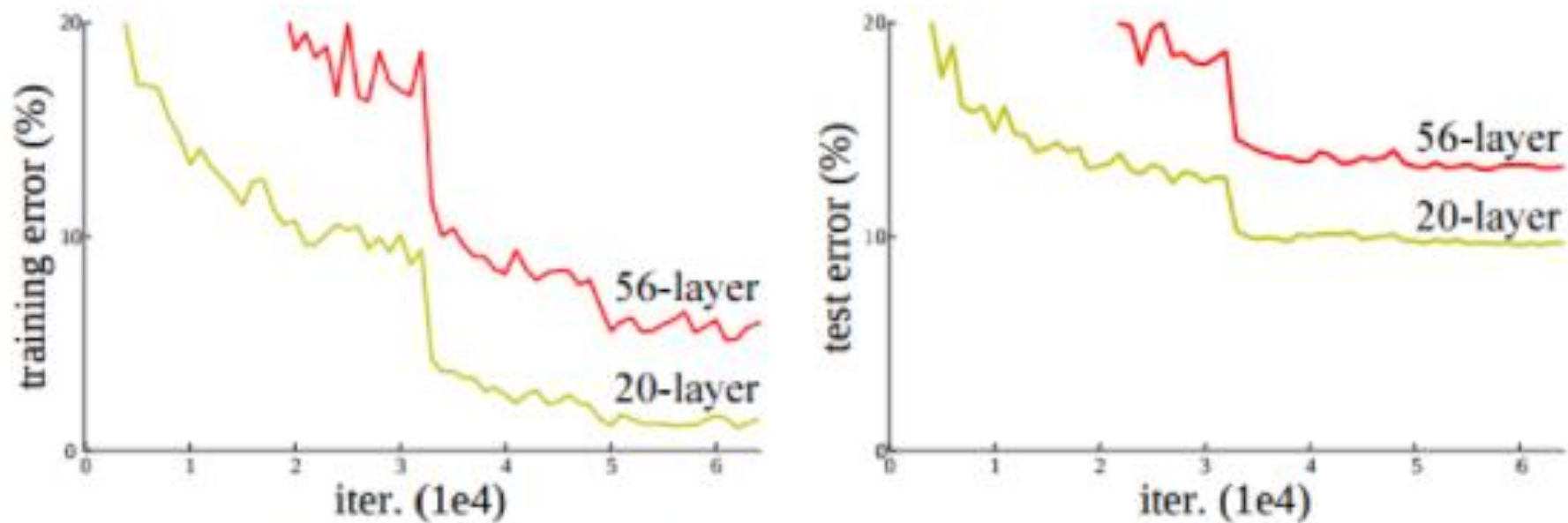


Figure 1. Training error (left) and test error (right) on CIFAR-10 with 20-layer and 56-layer “plain” networks. The deeper network has higher training error, and thus test error. Similar phenomena on ImageNet is presented in Fig. 4.

-> ResNet 팀이 망이 깊어지는 경우 어떤 결과가 나오는지 실험을 하였다.
20-layer 와 56-layer에 대해서 비교 테스트를 하였고, 결과는 위의 사진과 같다.

< Residual Learning >

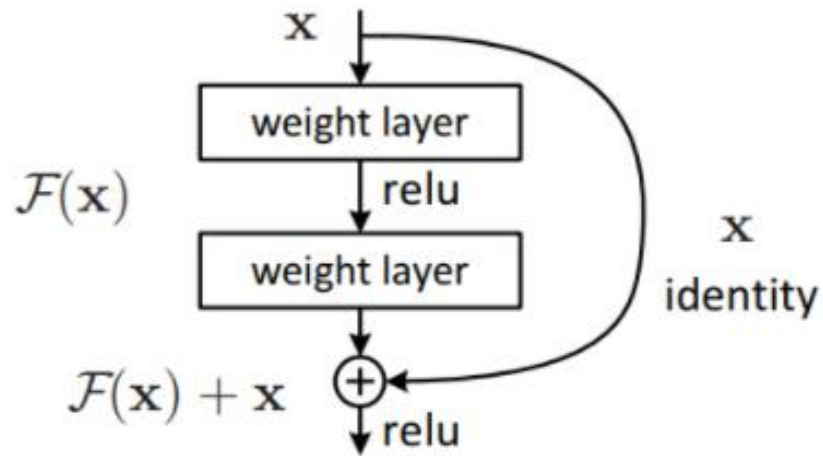
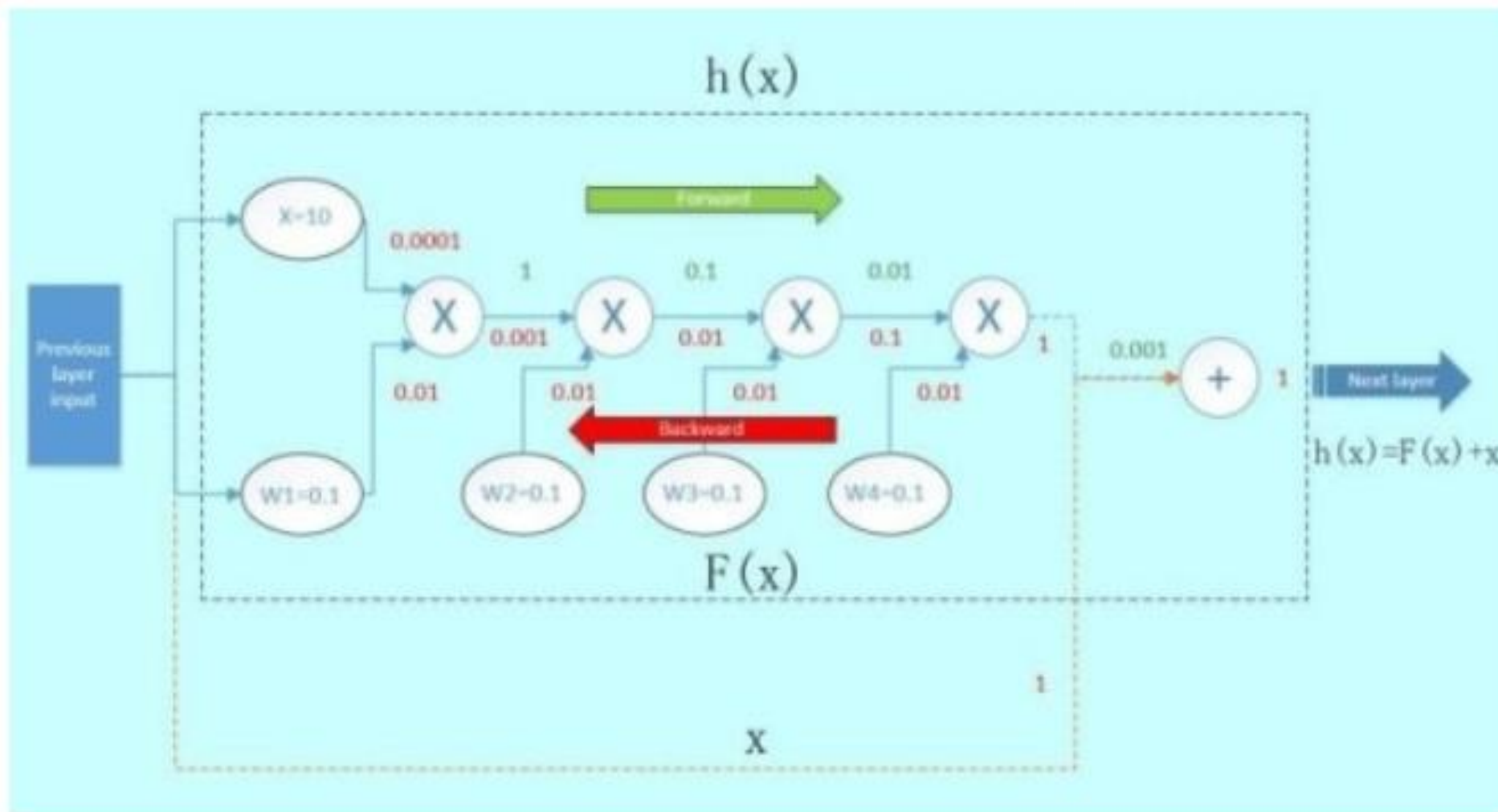


Figure 2. Residual learning: a building block.

- ResNet은 layer의 입력을 later의 출력에 바로 연결시키는 '**skip connection**'을 사용했다.
- > 출력 $H(x) = F(x) + x$ 로 변경이 되었다.
 - > 이러한 시도의 결과는 뛰어난 성능을 보여주었다.
 - > 기존 Network는 $H(x)$ 를 얻기 위한 학습을 하였지만 $F(x) = H(x) - x$
 - > $F(x)$ 가 0이 되는 방향으로 학습을 한다.
 - > x 가 그대로 skip connection이 되기 때문에 연산의 증가가 없다.
 - > fully - connected layer 뿐만 아니라 convolution layer에도 적용이 가능하다.
 - > x 와 $H(x)$ 의 dimension이 다르면 parameter w 를 추가해서 맞출 수 있다.

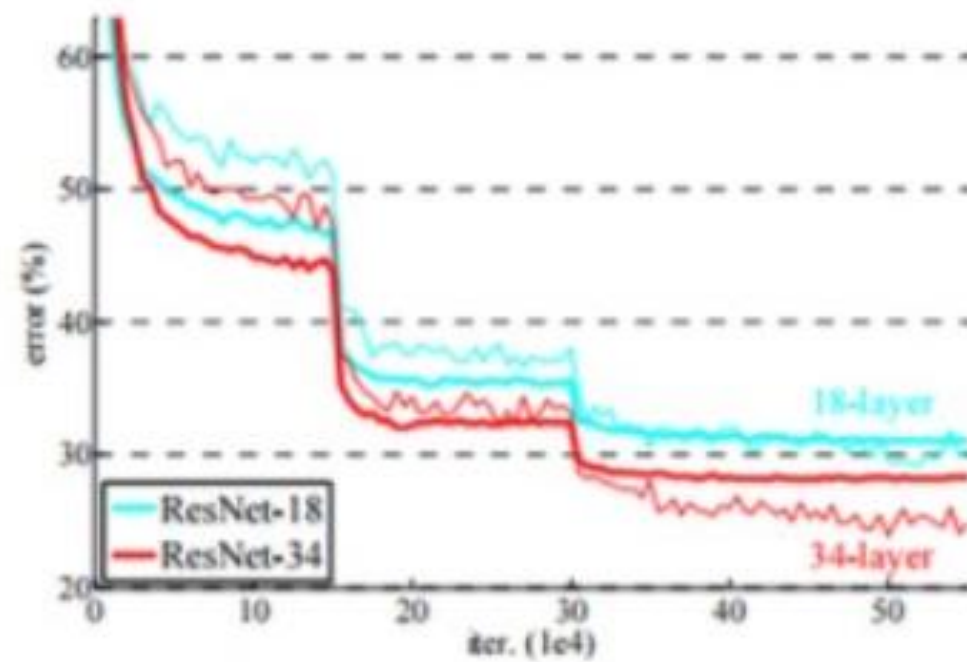
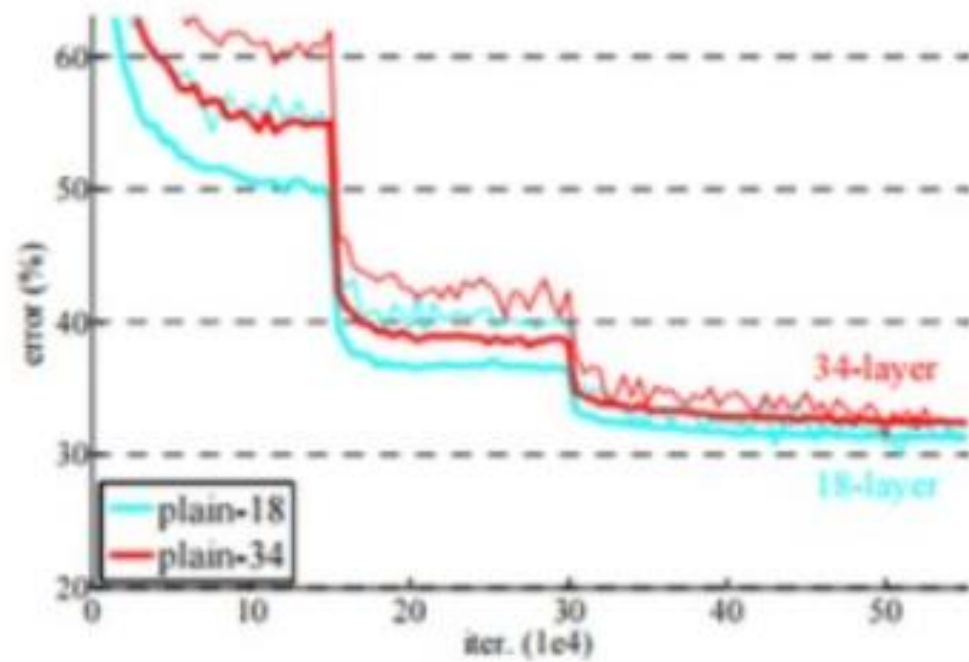


- ※ ResNet은 하나의 모듈에만 가능하다.
- ※ ResNet은 많은 모듈을 사용하며, 이 방식이 아니더라도 ReLU와 같은 활성화함수를 대체하는 것으로도 어느 정도 깊이의 Network까지는 gradient vanishing이 발생하지 않는다.

ResNet 팀은 성능 비교를 위해 일반적인 plain network와 비교를 하였다.



(실험 결과)



[개선된 ResNet 구조]

2015년에 발표한 ResNet은 $H(x) = \text{ReLU}(F(x) + x)$ 의 수식을 가지고 있습니다.

ReLU라는 activation function에 의해 실제로 blocking이 일어나게 됩니다.

다음 layer로 보내는 출력이 앞 layer에서 입력으로 가져온 온전한 skip connection이 아닙니다.

이러한 문제로 인해 ResNet팀은 ReLU를 $F(x)$ 안에 넣고,

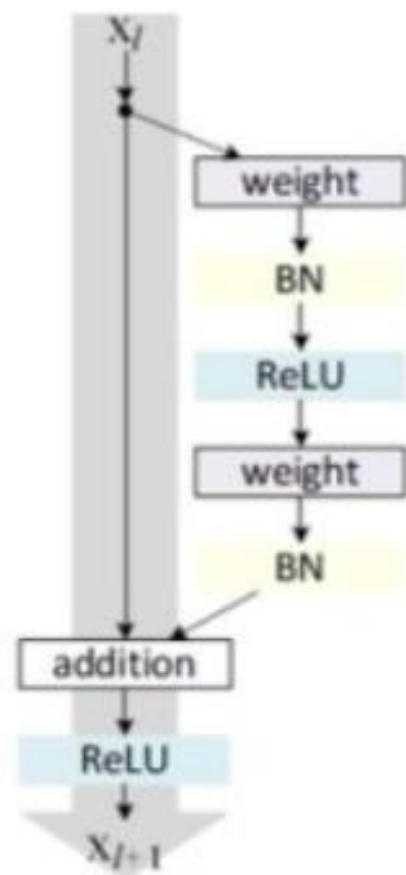
$H(x) = F(x) + x$ 라는 온전한 skip connection block을 설계하게 됩니다.

이걸 수식으로 표현하면 아래 그림과 같습니다.

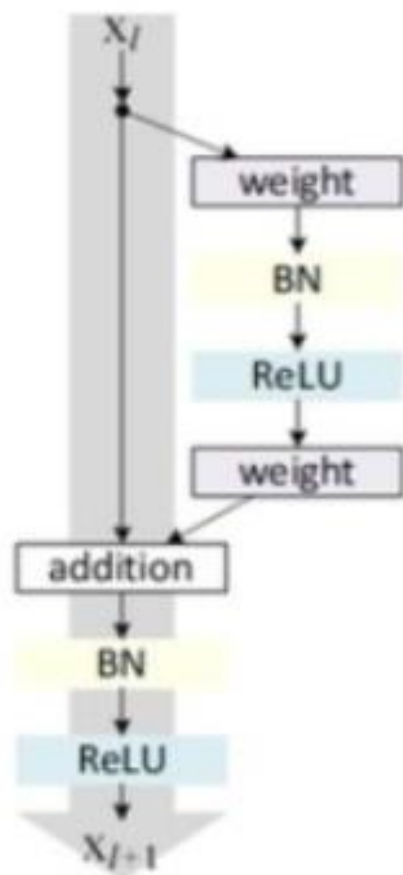
$$x_{l+1} = x_l + F(x_l) \quad \Rightarrow \quad \begin{aligned} x_{l+2} &= x_{l+1} + F(x_{l+1}) \\ x_{l+2} &= x_l + F(x_l) + F(x_{l+1}) \end{aligned}$$

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i)$$

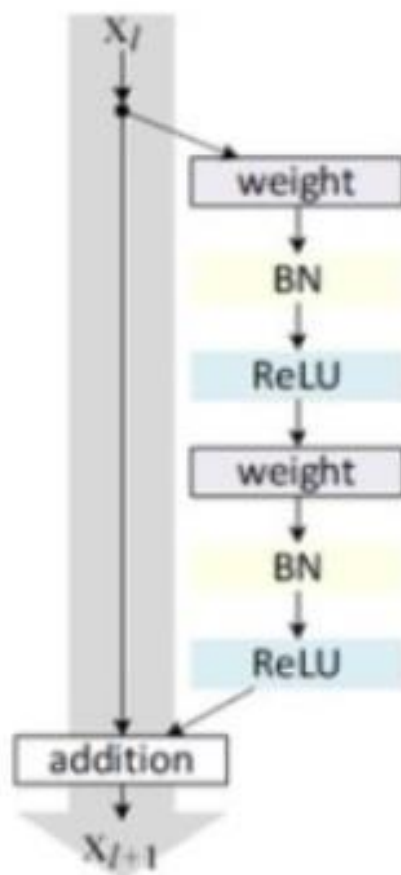
ResNet 팀은 activation 함수의 위치가 어떤 영향을 주는지 확인을 하기 위해 다음 그림과 같은 다양한 조합의 F(x) block에 대하여 실험을 진행하였습니다.



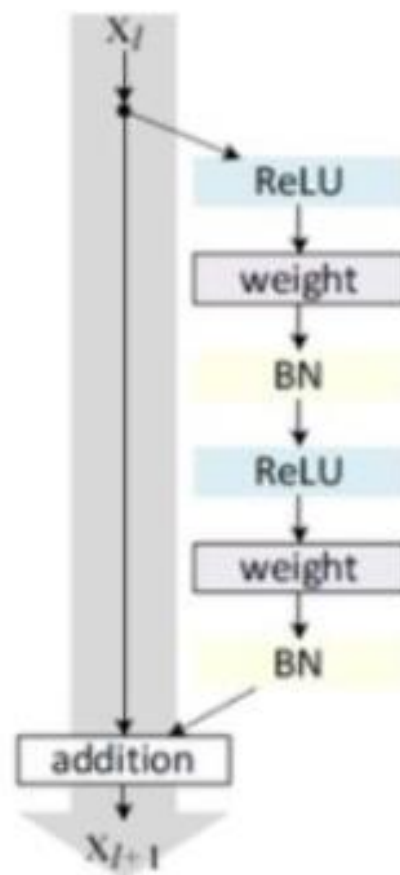
(a) original



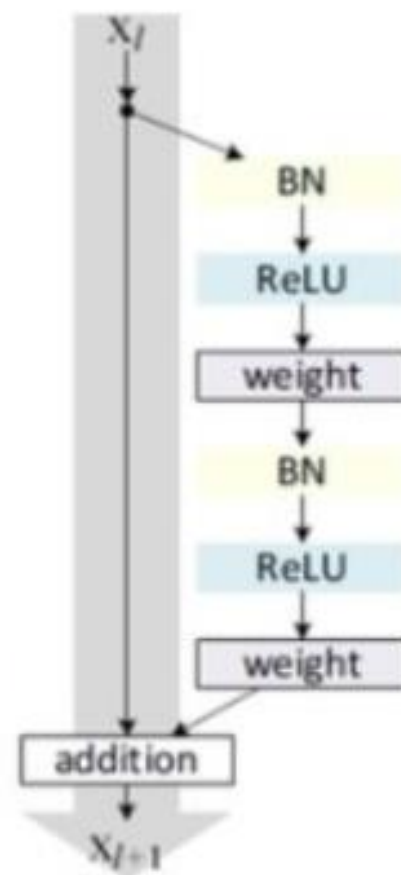
(b) BN after
addition



(c) ReLU before
addition



(d) ReLU-only
pre-activation



(e) full pre-activation

- (b)는 BN의 위치를 addition 뒤로 옮김.
결과적으로 BN이 propagation path를 방해할 수 있으며, origin보다 결과가 좋지 못하다.
- (c)는 addition 뒤에 있는 ReLU를 residual net쪽으로 옮김.
원래 residual 함수가 $(-\infty, +\infty)$ 범위인데, ReLU 때문에 결과가 non-negative가 나와서 representation 능력이 떨어지고 성능도 나빠진다.
- (d)와 (c)는 pre-activation 성능을 확인하기 위한 구조.
ReLU가 진정한 identity mapping 개념을 방해하기 때문에 이것을 움직이기 위한 테스트.

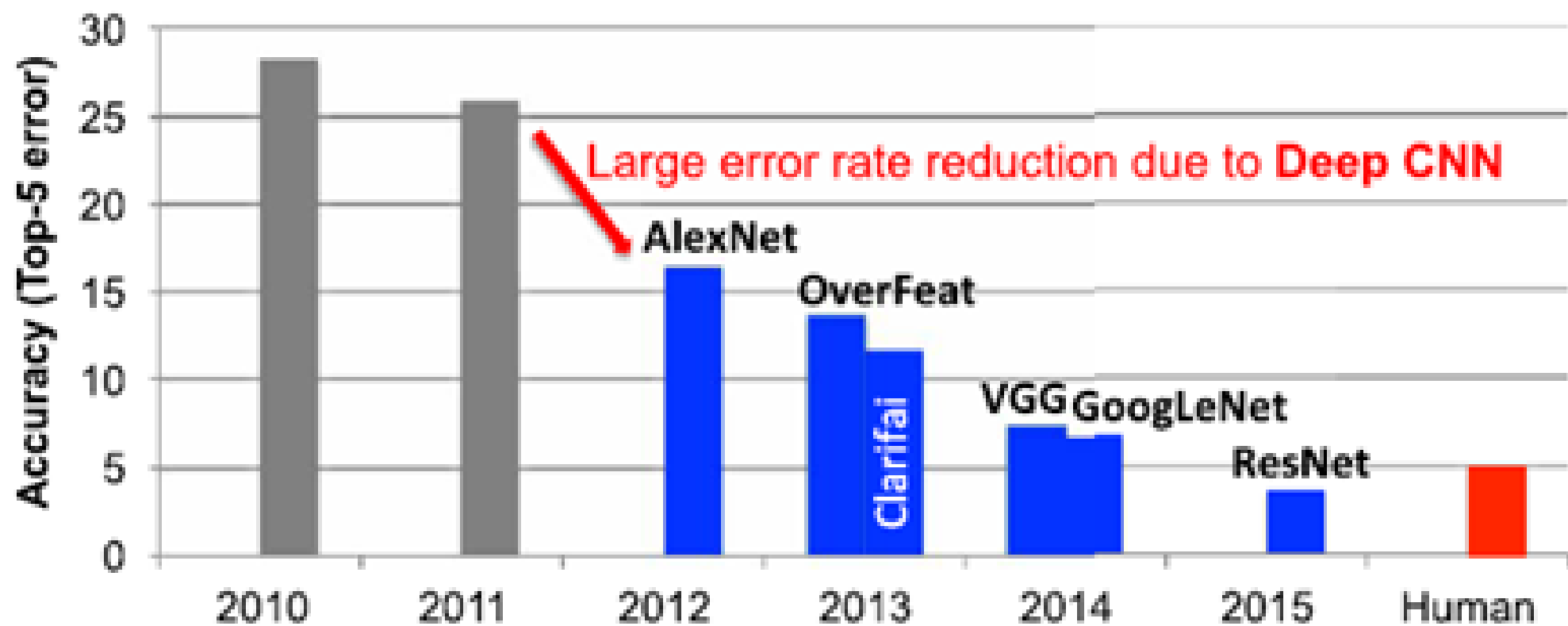


Fig. 7. Results from the ImageNet Challenge [14].

(Alex Net 구조) -> 2012 년 ILSVRC 대회 우승

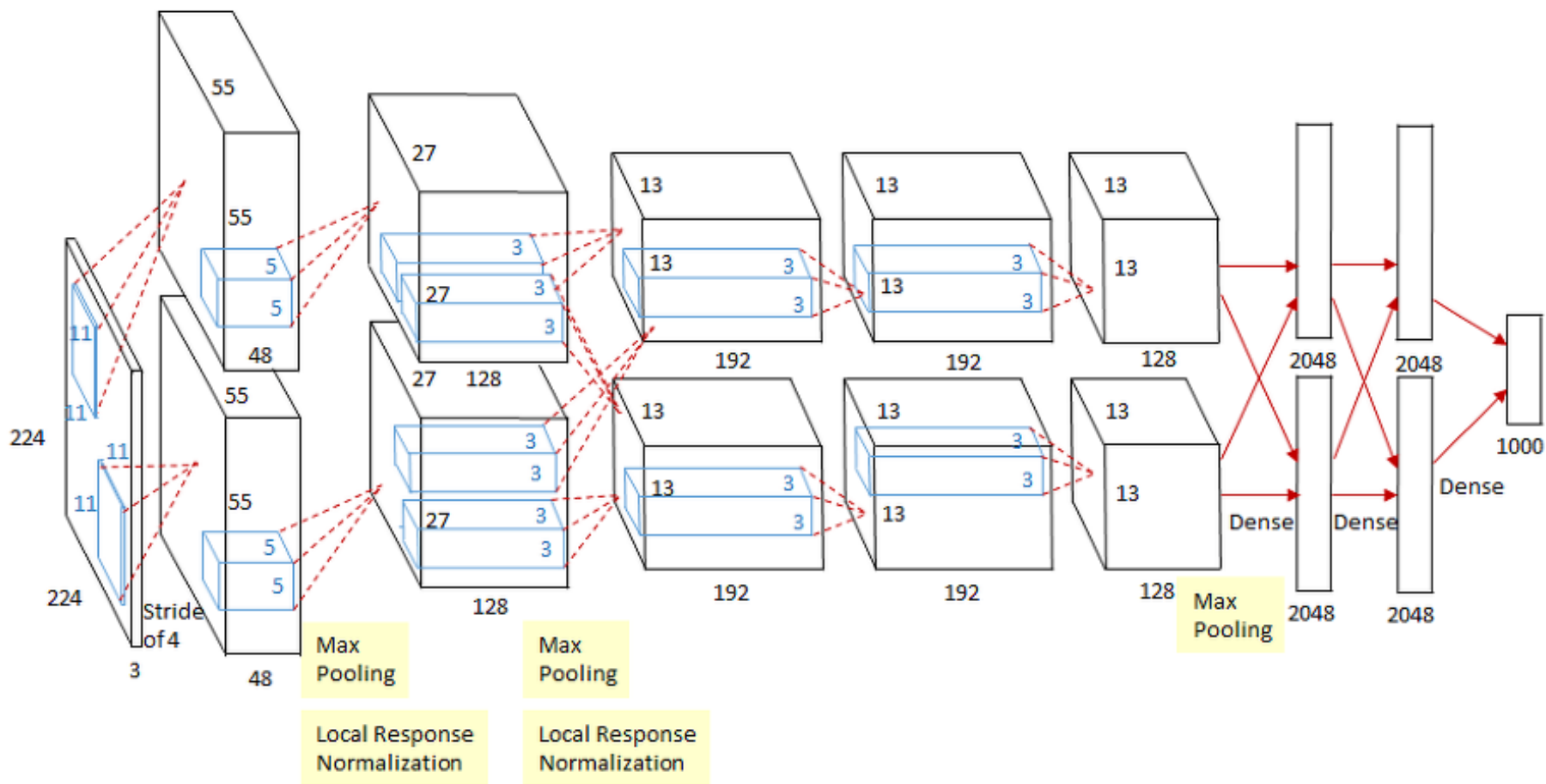
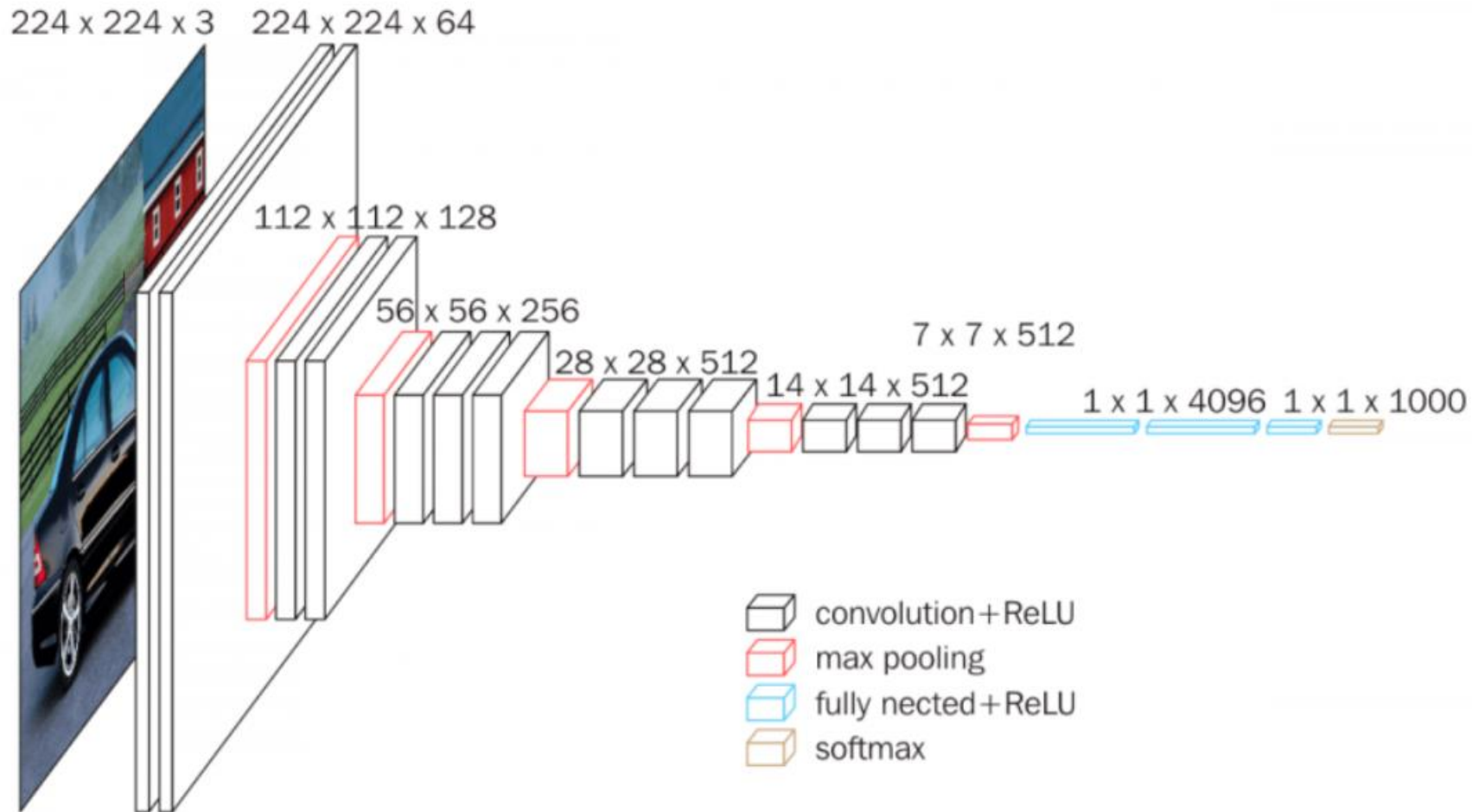


그림1. AlexNet의 구조도 [2]

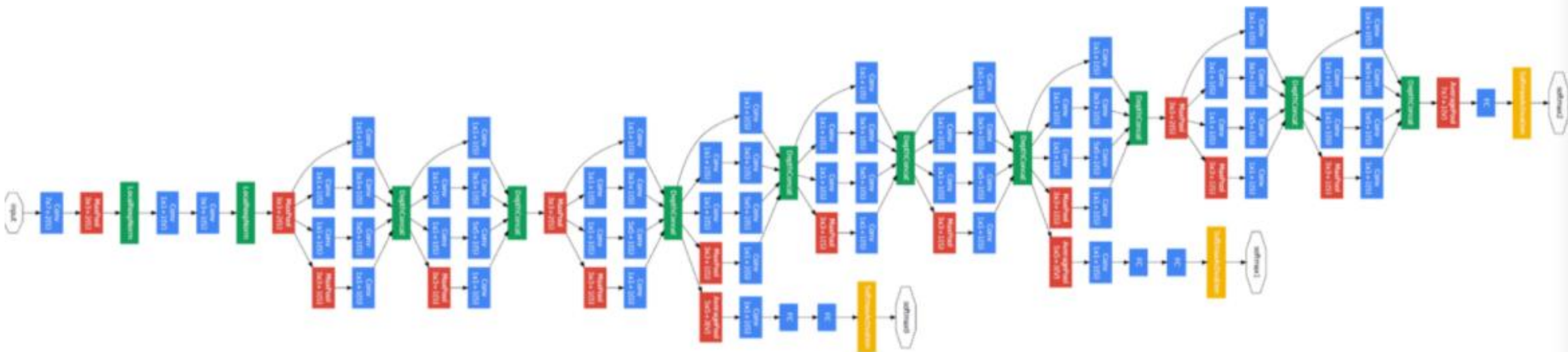
-> 8개의 layer로 구성되어 있다. (Conv layer : 5개, fully – connected layer : 3개)

(VGG-16 Net 구조) -> 2014 년 ILSVRC 대회 준우승



VGG16 구조 [6]

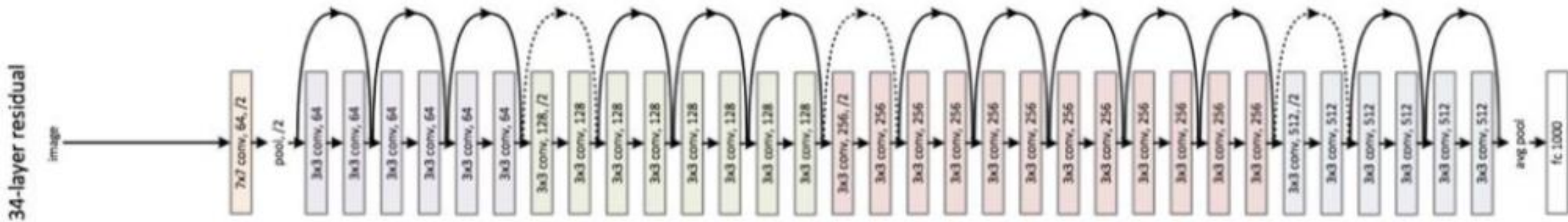
(GOOGLE Net 구조) -> 2014 년 ILSVRC 대회 우승



GoogLeNet Network (From Left to Right)

(Res Net 구조) -> 2015 년 ILSVRC 대회 우승

The great gradient highway.



< 프로젝트 진행 상황 >

현재 진행사항: Alex Net과 같이 여러 Convolution layer와 max pooling layer, Dropout layer, Fully connected layer 와 같은 층을 다양하게 쌓는 시도를 해보면서 Accuracy를 70~80% -> 90% 인근으로 향상시켰다.



남은 진행사항: training data set 을 추가 및 정제와 더불어 Res Net Model 과 같이 'skip connection' 기술을 사용하여 Accuracy를 95%이상으로 올리는 것이 남은 기간 목표이다.