

ResNet

Files	
Status	공부중
keyword	
part	딥러닝

블로그가 아주 잘 되어 있어서 일단 블로그만 첨부함...

8. CNN 구조 3 - VGGNet, ResNet

<https://m.blog.naver.com/PostView.nhn?blogId=laonple&logNo=221259295035&proxyReferer=https%3A%2F%2Fwww.google.com%2F>

Keywords....

- Skip connection

: input x 에 대해서 conv 1과 conv2를 거쳐서 연산된 결과를 $F(x)$ 라고 하자. (전통적인 CNN 방식이다)

$F(x)$ 에 x 를 더해준 $F(x)+x$ 를 최종 결과 $H(x)$ 라고 생각해보자. x 는 최종 결과인 $H(x)$ 를 계산하기 위해서 사용되었지만 conv1과 conv2를 건너뛰었다.

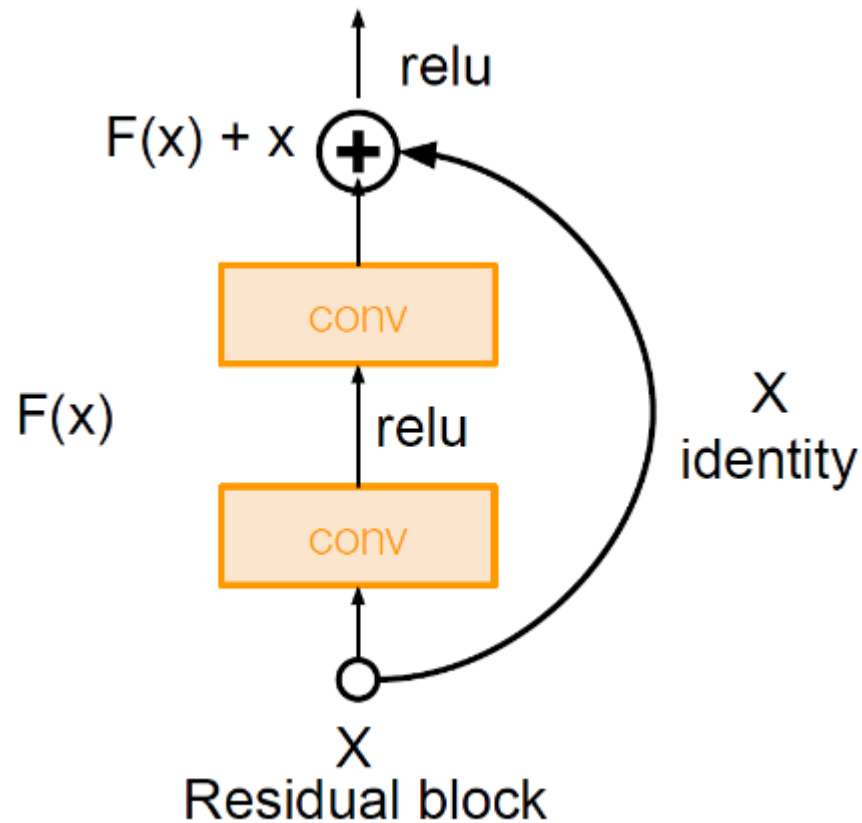
그래서 Skip connection이라고 한다.

- residual(나머지)

: $H(x)$ 를 학습한다는 말은 $F(x)+x$ 를 학습한다는 말과 같다.

$F(x)$ 를 학습한다는 말은 $H(x)-x$ 를 학습한다는 말과 같다.

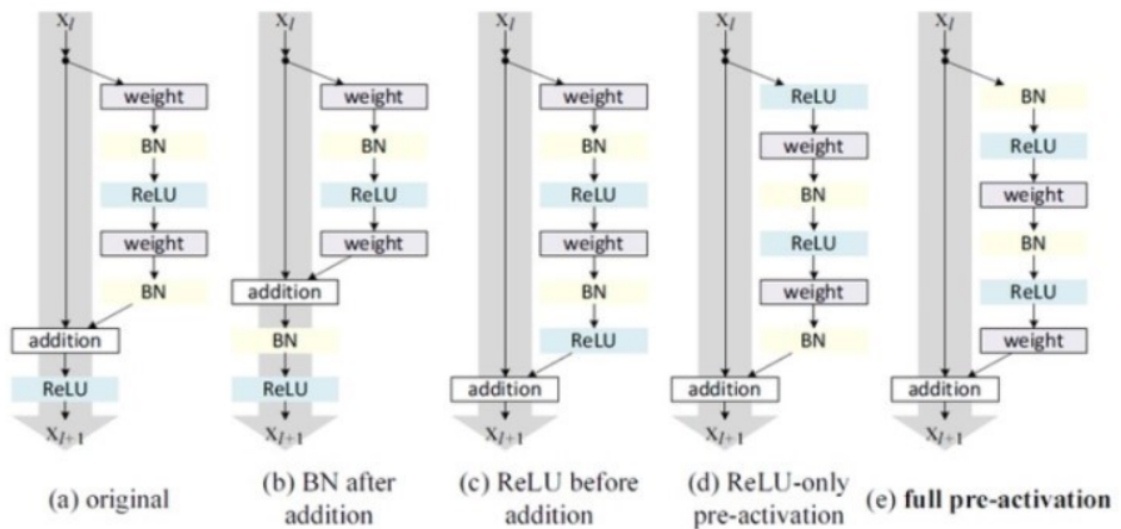
그래서 x 를 제외한 나머지(residual)을 학습한다는 뜻에서 ResNet이라는 이름이 붙었다.



- gradient vanishing 해결
 - x 를 계속 더해주므로 값이 0에 수렴하는 걸 방지할 수 있다.
- module화
 - skip connection은 module 단위로 적용된다.
 - module 내에서 layer는 얼마든지 쌓을 수 있다.
 - 단, gradient vanishing 문제는 module 단위로만 해결할 수 있다.
 - ResNet은 대개 여러개의 module을 쌓는 형태로 사용한다.



- Fully Connected Layer에서도 적용할 수 있고, Conv에도 적용할 수 있다.
 - Conv에 적용할 때 x 와 $F(x)$ 의 dimension이 안 맞으면 parameter w 를 추가하고 학습함.
- Layer가 깊을 수록 좋다.
- Activation Function을 붙이는 위치에 따라서 결과가 다르다.
 - 연구 결과가 있음. 어디에 붙이는게 제일 좋은지는 아직 찾아보지 않았음.



- Ensemble 효과

- ResNet은 어떤 Layer를 삭제한다고 해서 검사에 영향을 주지 않는다.
- 즉, 각 Layer는 독립적이며 상호 의존적이지 않다.
- Layer를 많이 삭제할 수록 Error는 높아진다.

이 페이지 url:

<https://www.notion.so/ResNet-aa40c471d78f4335aeb68a824576635a>