

## 1. Deep Residual Learning for Image Recognition

deep NN을 학습시키는 방법으로 residual learning framework를 고안함.

예전의 방식들보다 지속적으로 깊어져도 학습 가능.

learning residual function with reference를 이용함.

optimize하기 쉽고 depth가 깊어도 정확도를 얻을 수 있음.

imageNet dataset으로 residual net을 이용해 VGG보다 8배 깊은 network evaluate.

residual net이용해서 COCO object detection dataset으로 28% 성능 개선함.

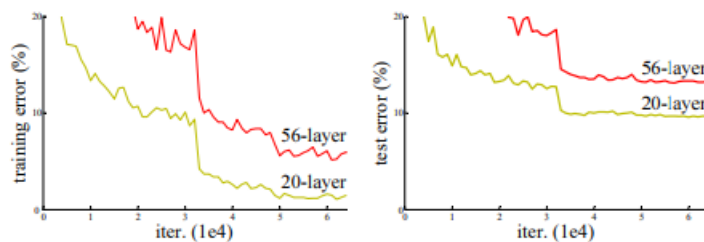


Figure 1. Training error (left) and test error (right) on CIFAR-10 with 20-layer and 56-layer “plain” networks. The deeper network has higher training error, and thus test error. Similar phenomena on ImageNet is presented in Fig. 4.

“degradation problem” = network depth 깊어질수록 accuracy가 saturated.

overfitting때문이 아니고 network 깊게 쌓을수록 training error가 높아져서 생기는 문제.

이러한 문제를 deep residual learning framework로 해결.

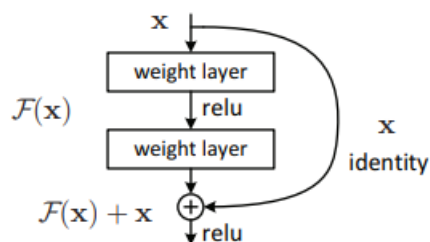


Figure 2. Residual learning: a building block.

Unreferenced mapping보다  $F(x)+x$ 를 optimize하는게 더 쉽다.

Shortcut connection = identity mapping 의 output을 stacked layer에 더함.

- + ) Extra parameter나 계산복잡도를 더하지 않는다. 그래서 좋다.
- + ) 전체 network는 여전히 end-to-end로 SGD이용해 backpropagation해서 train가능.
- + ) 라이브러리로 간단하게 구현 가능.

ImageNet dataset, CIFAR-10 dataset으로 residual 의 성능 보여줌.

- 1) 매우 깊은 residual net이 easy to optimize하다.
- 2) 이전의 network보다 깊어도 쉽게 accuracy 얻을 수 있다.

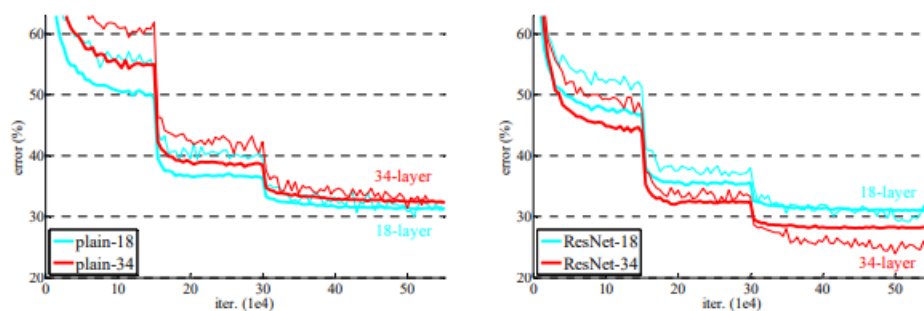


Figure 4. Training on **ImageNet**. Thin curves denote training error, and bold curves denote validation error of the center crops. Left: plain networks of 18 and 34 layers. Right: ResNets of 18 and 34 layers. In this plot, the residual networks have no extra parameter compared to their plain counterparts.

## 2. A Style-Based Generator Architecture for Generative Adversarial Network

Traditional GAN은 latent vector  $z$ 를 바로 input으로 넣어준 반면

StyleGAN에서는 학습된 constant값을 넣어준다.

그리고 mapping network와 noise가 추가된다.

Disentanglement : 각 변수가 하나의 변동 요인을 제어하는 linear subspace로 구성된 latent space.

Cf) latent vector (잠재벡터)란

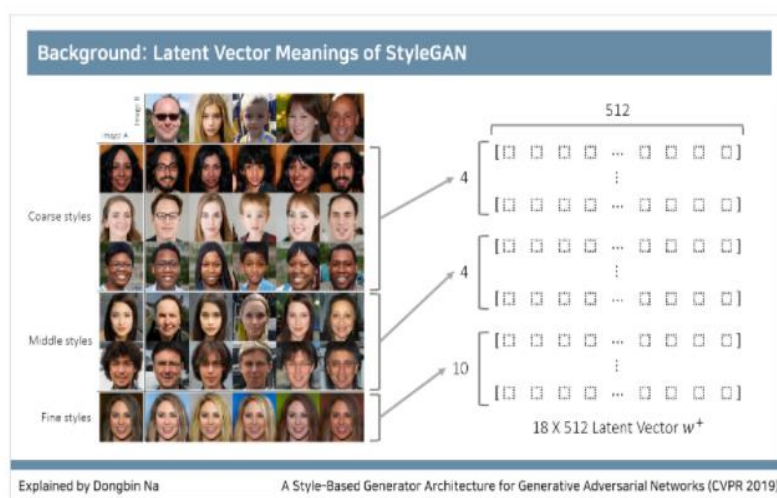
ex) 얼굴을 그리기 위해 성별과 바람의 방향을 잠재 변수에서 선택하는 경우를 생각해 봤을 때,

'여성이라면 바람의 방향은 높은 확률로 왼쪽이어야 한다'와 같은 두 잠재 변수 사이에 상관관계가 있다면 잠재 벡터 선택이 어렵다

의미 있는 잠재 벡터값을 쉽게 생성하기 위해서는 **잠재 벡터의 각 성분간 상관관계가 없도록** 학습이 되어야함.

**disentanglement** = 각각의 특징이 서로 얽혀있지 않고 적절히 풀려있다.

각 특징에 대한 분리가 가능하기 때문에 styleGAN 을 활용하면 editing 측면에서 굉장한 이점이 있음.



- styleGAN의 latent vector는 3가지 정도로 분류됨
  - coarse (거친) style
  - middle style
  - fine style

#### 1. Coarse(굵직한 특징) 저해상도

- 8 해상도까지 (\* 4x4 layer ~ 8x8 layer)- 포즈, 일반적인 헤어스타일, 얼굴형 등에 영향

#### 2. Middle(중간 특징) 중해상도

- 16부터 32 해상도까지 (\* 16x16 layer ~ 32x32 layer) - 자세한 얼굴 특징, 헤어스타일, 눈 뜨고/감음 등에 영향

#### 3. Fine(자세한 특징) 고해상도

- 64부터 1024(\* 64x64 layer ~ 1024x1024 layer) 해상도까지 - 눈, 머리, 피부 등의 색 조합과 미세한 특징 등에 영향

### 3. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising

네트워크 구조 디자인을 위해 vgg network를 image denosing에 맞게 변형,  
효과적인 path size를 기반으로 network의 depth 결정,  
Model 학습시에는 residual learning formulation을 사용.

### DnCNN Network Depth

3x3 conv filter, pooling layer 없음. Denosing neural network의 receptive field size는 effective path size와 관련. High noise level은 큰 effective patch size 요구.

### Architecture

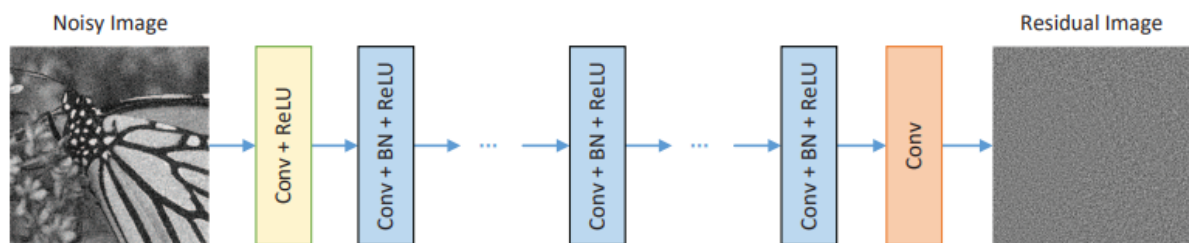
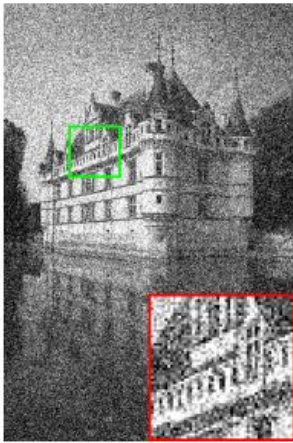


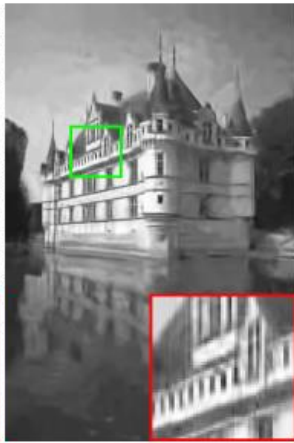
Fig. 1. The architecture of the proposed DnCNN network.

$R(y)$ 를 학습시키기 위해 residual learning formulation을 사용함.

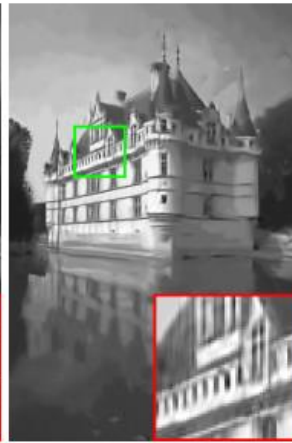
학습 속도와 denosing performance를 향상시키기위해 batch normalization도 함.



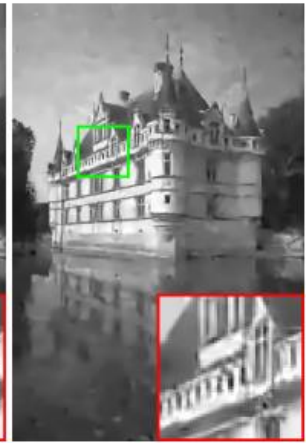
(a) Noisy / 14.76dB



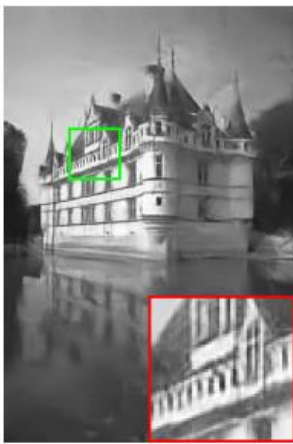
(b) BM3D / 26.21dB



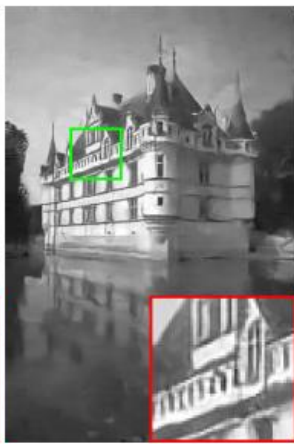
(c) WNNM / 26.51dB



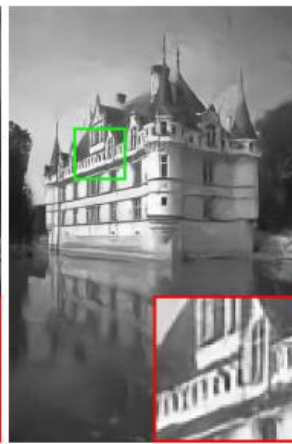
(d) EPLL / 26.36dB



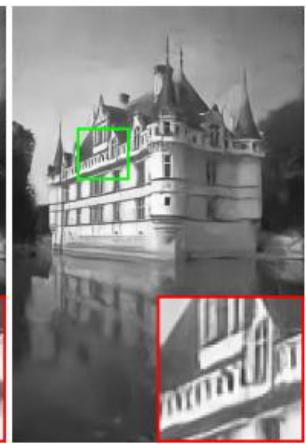
(e) MLP / 26.54dB



(f) TNRD / 26.59dB



(g) DnCNN-S / 26.90dB



(h) DnCNN-B / 26.92dB