

[논문리뷰] Beyond a Gaussian Denoiser : Residual Learning of Deep CNN for Image Denoising

Abstract

DnCNN 은 CNN을 이용하여 이미지에서 denoising 을 구현한다. 논문에서는 noise 의 예시로 Additive White Gaussian Noise(AWGN) 을 제거하고자 하였다. 기존의 방식들은 계산 시간이 너무 오래 걸리고, 파라미터 설정에 있어서 복잡한 인간의 개입이 필요한 문제가 있었다.

이 논문에서는 이미지를 직접 denoising 하는 것이 아니라 이미지에서 noise 를 분리해내는 것에 초점을 맞추었다.

또한 residual learning 과 batch normalization 을 사용하여 denoising performance를 높였다. 그리고 CNN을 사용했기 때문에 denoising 의 적용에 있어서 유연성을 확보할 수 있었다.

1. Introduction

Image denoising 은 오래됐지만 아직도 활발하게 연구되는 주제이다.

Denoising 에는 두 가지 큰 단점이 있는데

1. Those methods generally involve a complex optimization problem in the testing stage
2. The models in general are non-convex and involve several manually chosen parameters.

위 문제를 극복하기 위해 해당 논문에서는 아래와 같은 이유로 CNN을 사용한다.

1. CNN with very deep architecture is effective in increasing the capacity and flexibility for exploiting image characteristics.
2. Considerable advances have been achieved on regularization and learning methods for training CNN

이런 방법들은 CNN이 더 빨리 학습되게 하고 더 좋은 성능을 보이게 한다.

본 논문에서는 DnCNN 은 residual image 인 v 를 예측한다.

BN을 사용했을 때 DnCNN 의 성능과 안정성이 높아진다.

Contribution 은 크게 세가지이다.

1. Gaussian denoising 을 위한 end - to - end CNN을 제안한다.
2. Residual learning 과 BN을 사용하면 더 좋다.
3. Gaussian denoising 외의 JPEG deblocking 등의 다른 노이즈에 대해서도 성능이 뛰어나다.

2. Related Work

A. Deep Neural Networks for Image Denoising

Multi layer perceptioin, trainable nonlinear reaction diffusion, BM3D

B. Residual Learning

Residual Learning 은 원래 네트워크의 깊이가 깊어짐에 따라 생기는 성능 저하 문제를 해결하기 위해 제안되었다.

이로 인해 깊은 CNN도 쉽게 학습된다.

Residual network 들이 여러 residual unit 으로 이루어진 것과 달리 DnCNN은 residual unit 을 단 하나만 사용한다.

C. Batch Normalization

Mini-Batch SGD 는 CNN의 학습에 많이 사용되었지만, 이것의 simplicity 와 effectiveness 에도 불구하고 internal covariate shift 때문에 학습 성능이 떨어졌다.

3. The Proposed Denoising CNN Model

Network Depth

다른 연구를 참고해서 Conv filter 의 크기는 3x3 으로 고정했지만, pooling layer는 모두 지웠다. 그래서 깊이 d 인 모델의 receptive field 는 $(2d + 1) \times (2d + 1)$ 이 되어야한다.

DnCNN의 깊이를 정하기 위해 noise level을 시그마 = 25 로 고정하고 여러 denoising 방법에서의 효과적인 patch 크기를 조사했다.

아래 표에는 noise level 시그마 = 25 일 때 제일 효과적인 patch 크기를 나타낸다.

아래 표를 보면 EPLL에서 효과적인 patch 크기는 36x36 으로 가장 작았다.

흥미롭게도, DnCNN 은 EPLL 과 비슷한 크기인 35x35 로 했을 때 그 성능이 제일 좋았다.

그래서 gaussian denoising 을 위한 모델은 35x35 와 깊이는 17로 하고, 다른 일반적인 image denoising task 를 위한 모델은 더 큰 patch 를 사용하고 깊이는 20으로 했다.

TABLE I
THE EFFECTIVE PATCH SIZES OF DIFFERENT METHODS WITH NOISE LEVEL $\sigma = 25$.

Methods	BM3D [2]	WNNM [13]	EPLL [33]	MLP [24]	CSF [14]	TNRD [16]
Effective Patch Size	49×49	361×361	36×36	47×47	61×61	61×61

Network Architecture

DnCNN의 input 은 노이즈가 있는 $y = x + v$ 이다. DnCNN에서 residual mapping을 학습 시키기 위해 residual learning formulation $R(y) = v$ 를 사용했다.

MSE로 표현하면 다음과 같다.

$$\ell(\Theta) = \frac{1}{2N} \sum_{i=1}^N \|\mathcal{R}(y_i; \Theta) - (y_i - x_i)\|_F^2 \quad (1)$$

1) Deep architecture

깊이 D인 DnCNN은 아래 figure1 처럼 세 종류의 layer가 있다.

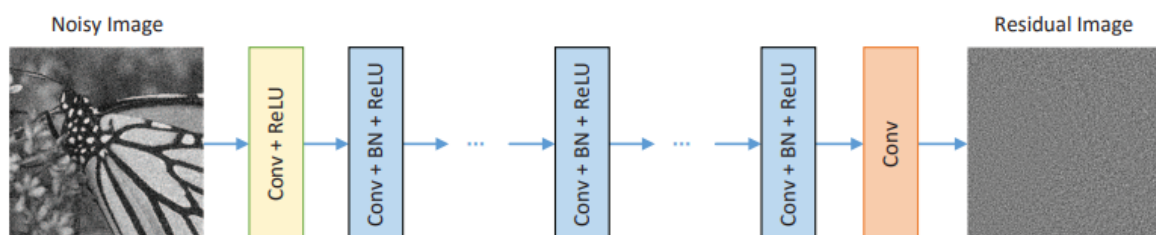


Fig. 1. The architecture of the proposed DnCNN network.

1. Conv + ReLU : 64 filters of size $3 \times 3 \times c$, (흑백일 때 $c = 1$, 색이 있을 때 $c = 3$)
2. Conv + BN + ReLU : $3 \times 3 \times 64$
3. Conv : $3 \times 3 \times 64$

2) Reducing Boundary Artifacts

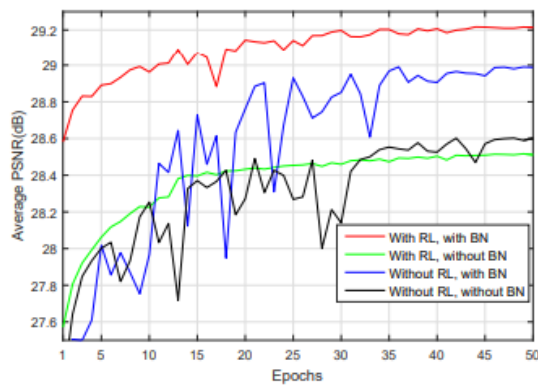
대부분 input size와 output size가 같기 때문에 boundary artifacts 가 생긴다.

이를 방지하기 위해 zero padding을 이용했다.

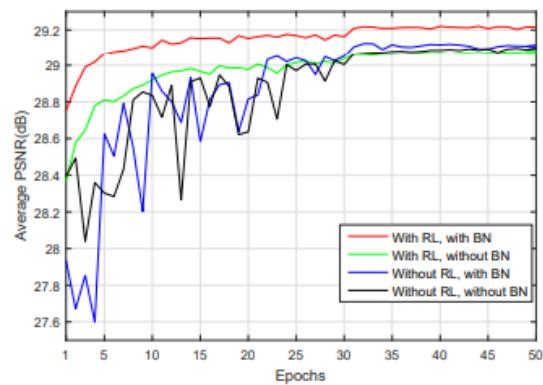
Integration of Residual Learning and Batch Normalization for Image Denoising

Figure 1 에서 original mapping $F(y)$ 가 x 를 예측하고, 동시에 residual mapping $R(y)$ 가 v 를 예측한다.

Original mapping 이 identity mapping 과 같을 때, residual mapping을 optimize하기 더 쉬워진다. Noisy한 y 는 residual 인 v 보다 clean 한 x 와 가깝기 때문에 $F(y)$ 가 $R(y)$ 보다 identity mapping에 가깝다.



(a) SGD



(b) Adam

위 그래프에서 residual learning 을 사용했을 때 더 빨리 안정적으로 converge 된다는 것과, RL 과 BN을 같이 사용했을 때 그 성능이 더 좋아진다는 것을 알 수 있다.

Connection with TNRD

TNRD는 다음 문제를 해결하기 위한 것이다.

$$\min_{\mathbf{x}} \Psi(\mathbf{y} - \mathbf{x}) + \lambda \sum_{k=1}^K \sum_{v=1}^N \rho_k((\mathbf{f}_k * \mathbf{x})_v),$$

위 식을 정리하면 아래와 같다.

$$\mathbf{v}_1 = \mathbf{y} - \mathbf{x}_1 = \alpha \lambda \sum_{k=1}^K (\tilde{\mathbf{f}}_k * \phi_k(\mathbf{f}_k * \mathbf{y})), \quad (4)$$

따라서 Fig 1의 CNN구조는 3가지 측면에서 발전된 TNRD라 할 수 있다.

1. influence function 을 ReLU로 바꿔 CNN학습이 더 잘되게 한다.
2. image 특성을 더 많이 담기 위해 CNN의 깊이를 늘렸다.
3. BN으로 학습이 더 잘되게 한다.

Extension to General Image Denoising

MLP, CSF 나 TNRD 와 같은 Gaussian denoising 방법들은 fixed noise level 에 대한 모델들이다. Unknown noise 에 대해서 Gaussian denoising 을 할 때 주로 사용하는 방법은 noise level 을 측정한 후 해당 noise level 에 맞게 학습 된 모델로 사용하는 것이다. 이런 방법들로 SISR 이나 JPEG denoising 과같은 non-Gaussian noise 에도 적용할 수 있다.